Mário A. Perini

Gramática descritiva do português

Há muito tempo pesquisadores e professores de língua portuguesa vêm demonstrando uma insatisfação com as deficiências da gramática normativa, apontando o crescente desinteresse dos alunos por este campo do ensino. Dessa forma, as maiores lacunas dizem respeito à falta de coerência interna, à sua inadequação às realidades de nossa língua hoje e à má formulação de normas e definições.

Assim, apoiando-se nas contribuições e evoluções dos estudos lingüísticos dos últimos anos, Mário A. Perini elabora uma nova gramática e pretende com isso suscitar a discussão sobre os fatos que se observam num exame crítico da língua portuguesa atual, além da necessidade de se repensar o ensino de gramática.

Não é também objetivo do Autor instituir uma nova

MÁRIO A. PERINI

GRAMÁTICA DESCRITIVA DO PORTUGUÊS

Sumário

SEGUNDA PARTE: SINTAXE

Terceira Parte: Semântica

Quadros

Nota do Editor

Alguns leitores poderão achar que a linguagem desta *Gramática* se afasta do padrão estrito usual neste tipo de livro. Assim, o Autor escreve *tenho que reformular*, e não *tenho de reformular*, *pode-se colocar dois constituintes*, e não *podem-se colocar dois constituintes*, e assim por diante. Isso foi feito de caso pensado, com a preocupação de aproximar a linguagem da gramática do padrão atual brasileiro presente nos textos técnicos e jornalísticos de nossa época.

Aqui, como sempre, manifesta-se a preocupação em reconhecer um padrão existente nos textos escritos modernos, ao contrário de um padrão (que para o Autor é fictício) estabelecido pelas normas gramaticais.

Lista das principais abreviaturas utilizadas

AA	adjunto adverbial (função)
AC	adjunto circunstancial (função)
Ant	anteposição (traço)
AO	adjunto oracional (função)
Atr	atributo (função)
Cl	cliticizável (traço)
CN	concordância nominal (traço)
CP	complemento do predicado (função)
cr	co-referente (traço)
CSA	complemento do sintagma adjetivo (função)
CV	concordância verbal (traço)
Det	determinante (função)
Ex	(nas transitividades) exigência
Int	intensificador (função)
L	(nas transitividades) aceitação livre
ModE	modificador externo (função)
ModI	modificador interno (função)
NdP	núcleo do predicado (função)
NSA	núcleo do sintagma adjetivo (função)
NSN	núcleo do sintagma nominal (função)
Num	numerador (função)
NV	negação verbal (função)
PA	posição do auxiliar (traço)
PDet	predeterminante (função)
PNE	pré-núcleo externo (função)
PNI	pré-núcleo interno (função)
pNdP	posição obrigatória antes do NdP (traço)
Poss	possessivo (função)
Pred	predicado (função)
Pv	predicativo (função)
Q	retomável através de (*o*) *que/quem* (traço)
Qf	quantificador (função)
Rec	(nas transitividades) recusa
Ref	reforço (função)
SAdj	sintagma adjetivo
SAdv	sintagma adverbial
SN	sintagma nominal

Prefácio

insatisfação com a qualidade das gramáticas vem de longe e atinge todos os grupos interessados: lingüistas teóricos, especialistas e professores de português, alunos e simples curiosos. Depois de décadas apontando os defeitos da tradição gramatical, cobra-se insistentemente dos pesquisadores a elaboração de gramáticas mais adequadas.

Algumas tentativas já foram feitas nesse sentido: mencionarei Back & Matos, 1972; Macambira, 1982; e Mateus et al., 1983. A presente *Gramática* representa uma nova proposta, que espero seja melhor do que suas predecessoras. A seguir, farei algumas considerações sobre as condições do ensino gramatical no Brasil, para situar este livro no contexto em que ele pretende atuar.

A lingüística se tem desenvolvido grandemente nos últimos tempos; no Brasil passamos do quase nada da década de 60 até uma comunidade numerosa, com produção intensa, freqüentemente de boa qualidade, em praticamente todas as grandes áreas da disciplina. Hoje se faz lingüística de bom nível entre nós; lançam-se as bases para uma descrição coerente, empiricamente adequada, teoricamente sofisticada de todos os aspectos da língua, de seu uso, variação, aquisição, evolução histórica e assim por diante. Descobrem-se coisas novas e reinterpretam-se velhas descobertas, lançando mais luz sobre a nossa língua, sobre a realidade lingüística do país e sobre a linguagem em geral.

Enquanto tudo isso acontece, porém, nossos filhos são obrigados a estudar compêndios de gramática onde se considera a homonímia e a polissemia "defeitos da língua", e não recursos essenciais de comunicação; onde se definem funções e classes de maneira totalmente vaga e confusa; onde se fa-

la de um fonema como sendo um "som"; onde se justifica uma análise com base em construções desusadas desde o século XIX; e assim por diante.

Esse não é evidentemente um problema lingüístico. Mas a lingüística é feita por pessoas, e a situação do ensino gramatical, como a do ensino de língua portuguesa em geral, afeta profundamente o desenvolvimento dos estudos lingüísticos (para não falar do desenvolvimento intelectual dos jovens).

Para avaliar a gravidade da situação, basta perguntar a alguns alunos de segundo grau se eles contemplariam dedicar sua vida aos estudos gramaticais: dificilmente se obterá resposta afirmativa. Na melhor das hipóteses, estão sendo formados jovens cuja curiosidade intelectual é dirigida em qualquer direção, menos na dos estudos da língua. Na pior das hipóteses, estão sendo formados jovens cuja aversão aos estudos gramaticais os leva a sufocar a própria curiosidade intelectual.

Existe um lugar para a gramática no ensino de primeiro e segundo graus, lugar atualmente ocupado por uma doutrina gramatical que prescinde de toda a lingüística dos últimos setenta anos — e, na verdade, também de boa parte da lingüística anterior. Trata-se da gramática tradicional; ou, melhor dizendo, da versão atual, seriamente empobrecida, da velha gramática tradicional.

Do jeito como estão as coisas, os estudos gramaticais têm feito mais mal que bem. São o último reduto do *magister dixit,* da doutrina oficial, no mundo moderno. O que se tem feito a respeito? Alguma coisa, evidentemente, mas sempre algo fragmentária, por falta de uma descrição abrangente. Em meus cursos e palestras para professores, detecto sempre uma insatisfação generalizada; mas também um grande sentimento de impotência, por falta de alternativas viáveis.

Sustento a idéia de que é responsabilidade dos lingüistas brasileiros elaborar uma descrição ampla e detalhada da realidade lingüística do Brasil. Entendo essa descrição como abrangendo não só a estrutura da língua padrão (objeto da minha *Gramática*), mas ainda a descrição da língua coloquial, sua variação social e geográfica, sua história etc.

Há muita coisa feita, mas totalmente inacessível aos que trabalham na escola. Alguém precisa parar um pouco e tentar produzir textos que alcancem, já não digo o aluno, mas pelo menos o professor. É preciso mostrar que há alternativas, e melhores do que o que está aí. É preciso, acima de

tudo, mostrar que é possível, e necessário, refletir sobre os problemas da linguagem.

Construir uma nova gramática é uma humilde parte dessa tarefa. Não tenho ilusões quanto à importância relativamente reduzida dos estudos gramaticais *stricto sensu*, frente ao estudo de outros aspectos da língua (ver, a respeito, a seção 1.3). Mas a construção da gramática é um dos passos que precisam ser dados, e isso lhe confere importância suficiente.

Uma gramática é obra de síntese e, como tal, depende do estado da pesquisa em cada uma das áreas consideradas. O resultado, portanto, é desigual: em certos pontos, é possível incluir na descrição resultados de pesquisas relevantes e razoavelmente completas; em outros casos, é preciso contentar-se com indicações mais programáticas; às vezes, nada mais se pode fazer do que definir o problema e clamar por maiores pesquisas.

Nesta *Gramática*, apresentei uma série de problemas, tentando formulá-los com a clareza possível; e para a maioria deles apresentei soluções. Não tentei evitar soluções controversas e muitas vezes apontei as fraquezas da análise adotada. Espero, evidentemente, que muitas das soluções aventadas possam ser aceitas como corretas; mas não é isso o essencial. O mais importante é suscitar discussão — discussão de problemas reais da análise da língua, baseada em fatos observáveis do português padrão atual. Meu objetivo é contribuir para uma reorientação radical dos estudos de língua portuguesa: do exame das gramáticas para o exame dos fatos da língua — sem desprezar o exame das gramáticas, mas submetendo-as a uma crítica rigorosa.

Os estudos de gramática portuguesa tendem atualmente a reduzir-se ao exame da literatura anterior (que, por sua vez, muitas vezes se limita a repetir ou parafrasear a literatura ainda mais antiga), complementado, ocasionalmente, com opiniões muito pouco justificadas. Observa-se em alguns casos uma tentativa, sempre muito tímida, de lançar mão de dados da língua atual (como quando se admite a existência de construções do tipo *ele vai me telefonar*).

É bem verdade que faltam estudos empíricos suficientemente amplos e confiáveis nos quais basear uma análise do padrão moderno brasileiro. Mas há pontos que jamais são questionados, e que poderiam sê-lo desde já, como as próprias bases teóricas da gramática e a atitude normativista que atualmente a permeia. Este livro pretende ser um exemplo de como se pode fazer uma gramática portuguesa sobre bases teoricamente coerentes, ao mesmo

tempo isolando as preocupações normativas (que, em si, não são necessariamente condenáveis) da descrição da língua propriamente dita.

Se o mais que se conseguir for suscitar críticas, debates e discussões, o objetivo principal terá sido alcançado.

No que pesem as preocupações de ordem escolar que motivaram em parte este trabalho, a presente *Gramática* não tem a finalidade de ser utilizada nas salas de aula de primeiro ou segundo graus. Ao preparar o texto, pareceu-me indicado dirigir-me, em um primeiro momento, aos profissionais da área: professores de primeiro e segundo graus, alunos e professores dos cursos de Letras. Como se verá, há muita novidade a ser apreendida, assimilada e — principalmente — discutida; só depois será possível pensar em sua aplicação ao ensino de primeiro e segundo graus.

Muitos leitores estranharão o tom bastante "científico" desta *Gramática*; poderão ter alguma dificuldade na leitura de passagens onde as análises não são apenas expostas, mas discutidas e nem sempre resolvidas de modo definitivo. Estou consciente de que o texto exige do leitor muito mais do que as gramáticas usuais — não apenas por requerer a assimilação de noções e termos novos, mas principalmente porque requer a formação de uma atitude nova frente ao estudo da gramática. Mas é este justamente o grande passo que, acredito, deve ser dado no desenvolvimento de nossa disciplina (ver a respeito a seção 1.3.). Não há razão alguma para que uma gramática seja menos "científica" em sua concepção e em sua redação do que um compêndio de biologia ou de psicologia. E não há razão para que o professor de gramática seja dispensado da formação científica que se exige de um professor de biologia ou de psicologia.

Em duas palavras, procuro ajudar o leitor a dar um passo que é inevitável e já vem muito atrasado. Creio que a gramática é a última disciplina a enfrentar a mudança. Até quando ficaremos estagnados em um modo de pensar ultrapassado e francamente contraproducente do ponto de vista educacional?

Assim, esta *Gramática* deve ser entendida como uma contribuição para a reformulação do ensino gramatical entre nós. Não é o passo final, evidentemente. É minha intenção produzir material para uso em sala de aula, baseando-me nos princípios aqui expostos. Mas é claro que esse material ficaria sem função sem um corpo importante de docentes suficientemente informados e formados para utilizá-lo. Acredito que após algum tempo de di-

vulgação das idéias contidas neste livro será possível estudar a possibilidade de incluí-las em livros didáticos de primeiro e segundo graus.

O livro se divide em quatro partes principais: na primeira discuto os objetivos do estudo da gramática (em especial na escola de primeiro e segundo graus) e exponho uma série de noções indispensáveis a quem deseja estudar seriamente a gramática do português. A segunda parte é dedicada à sintaxe; a terceira, à semântica; e a quarta a um exame muito sumário de alguns aspectos da lexicologia (privilegiando as classes de palavras).

O leitor notará uma ausência importante: a morfologia. Infelizmente, restrições de tempo me impediram de incluir uma parte que tratasse desse componente da gramática; pretendo sanar a lacuna em edições futuras.

No final do volume incluí alguns "quadros", onde são discutidos com um pouco mais de detalhe certos problemas suscitados pela análise; também aqui o tempo disponível me obrigou a deixar de lado vários temas interessantes. Nos quadros permiti-me ser um pouco mais técnico do que no restante do livro, desenvolvendo alguns assuntos em maior profundidade.

O objetivo principal do livro não é tanto dar uma noção completa do que se sabe sobre a estrutura do português; é, antes, discutir pontos relativamente bem compreendidos e ao mesmo tempo olhar para frente, identificando problemas a estudar. Por isso me preocupei em apontar temas de pesquisa futura, sempre que oportuno. É definitivamente necessário começar a conceber a gramática como uma disciplina viva, em revisão e elaboração constante.

Finalmente, alguns agradecimentos.

Dentre as pessoas que deram sua contribuição à feitura desta *Gramática*, algumas merecem citação especial. Em primeiro lugar, agradeço aos membros da equipe do projeto TENPo: Regina Bessa, Sigrid Fraiha e Lúcia Fulgêncio, cujo trabalho sobre a estrutura interna do sintagma nominal é responsável por algumas das novidades deste livro.

Agradeço também a Ev'Ângela de Barros, Jânia Ramos, Ataliba de Castilho, Giampaolo Salvi e Yara Liberato, que contribuíram com discussões, palpites e encorajamento. E gostaria de mencionar a contribuição de um leitor anônimo da Editora Ática, cujas observações foram muito úteis, especialmente no esclarecimento de passagens obscuras. Nem preciso dizer que nenhum desses amigos se compromete com a qualidade do trabalho, que é de minha exclusiva responsabilidade.

A pesquisa em que parte desta análise se baseia recebeu o auxílio de uma bolsa de pesquisa do CNPq; e, naturalmente, pude usar de maneira bastante livre meu tempo como professor da Universidade Federal de Minas Gerais. A essas instituições vai também o meu muito obrigado.

PRIMEIRA PARTE:
PRELIMINARES

1 Introdução

1.1. FISIONOMIA DESTE LIVRO

Esta *Gramática* tem objetivos muito pouco modestos: propõe uma maneira realmente nova de descrever a estrutura do português — partindo de princípios teóricos muito mais rigorosos do que aqueles em que se baseiam as gramáticas atualmente disponíveis, para chegar a uma análise bastante diferente da usual. Portanto, é um livro cheio de novidades, mesmo para quem vem estudando gramática portuguesa intensivamente há anos (talvez **principalmente** para esses).

As novidades aqui encontradas só são realmente novas enquanto aplicação a uma descrição abrangente da estrutura do português. Todos os princípios gerais e a maioria das noções particulares utilizadas nesta descrição têm sido lugar comum da pesquisa lingüística nas últimas décadas. O que este livro tem de novo é a tentativa de apresentar a gramática do português seguindo estritamente uma orientação teórica coerente e mais atualizada. Não se trata da primeira proposta nesse sentido; não é mais que justo mencionar aqui trabalhos como os de Back & Matos, 1972; Macambira, 1982; e Mateus et al., 1983, que representam tentativas análogas à do presente trabalho. No entanto, acho que essas análises deixam a desejar em muitos pontos, o que justifica a publicação de mais uma proposta.

Os estudos de gramática portuguesa estão seriamente defasados, de dois pontos de vista. Em primeiro lugar, têm sido influenciados por uma atitude questionável frente ao objeto de estudo e ao seu ensino. Discuti este problema em meu livro *Para uma nova gramática do português* (Ática, 1985), onde tentei identificar seus principais ingredientes, a saber: falta de coerência teórica, falta de adequação à realidade da língua e normativismo sem controle. Vou comentar rapidamente cada um desses pontos.

A falta de coerência teórica se manifesta, por exemplo, nas muitas definições que *não podem* ser segui-

das se se deseja identificar as entidades que elas pretendem definir. Um dos exemplos dados no livro mencionado é a definição de *sujeito* como "o termo do qual se afirma alguma coisa"; é bem fácil verificar que os termos usualmente analisados como sujeito freqüentemente não exprimem o ser do qual se afirma alguma coisa. No entanto, continua mantendo-se tanto a definição quanto a análise, muito embora as duas estejam em contradição.

A falta de adequação à realidade da língua aparece quando a gramática descreve (ou "recomenda") verdadeiras ficções lingüísticas: construções que caíram de moda há séculos, ou mesmo que jamais existiram. Um exemplo é a afirmação de que só se coloca um pronome clítico (oblíquo átono) entre um auxiliar e o verbo principal ligando-o ao auxiliar por ênclise, isto é, *estou-me divorciando* e não *estou me divorciando*. Ora, sabemos que, apesar da opinião dos gramáticos, a segunda forma é a mais comum na língua atual.

Quanto ao normativismo sem controle, terei algo a dizer a respeito mais adiante, na seção 1.3.4.

Outro aspecto da desatualização dos estudos gramaticais é a falta de incorporação dos resultados teóricos e práticos da pesquisa lingüística das últimas décadas. Não é exagero afirmar que as gramáticas portuguesas de hoje representam a situação dos estudos lingüísticos por volta dos princípios do século XX — quase um século de atraso, portanto. Aconteceu muita coisa nesse tempo, e muito pouco foi assimilado; como resultado, as gramáticas, além de todos os seus outros defeitos, são também arcaicas, tanto na descrição que oferecem quanto nas teorias em que se baseiam.

Este livro é também uma tentativa de enfrentar este último problema. Procuro apresentar uma análise que leve em conta os resultados obtidos nos últimos tempos pela pesquisa na área de gramática portuguesa. Minha preocupação inicial foi manter-me tão próximo quanto possível da análise tradicional, para não exigir do leitor um esforço maior do que o estritamente necessário. Apesar disso, como se verá, a análise ficou muito diferente da tradicional; acontece que sete ou oito décadas de atraso não podem ser transpostas com um pequeno passo. O grande desafio, em última análise, é encontrar uma linguagem capaz de exprimir os resultados da investigação lingüística das últimas décadas e que não exija para sua compreensão um treinamento prévio que, sei bem, a imensa maioria dos professores e alunos não possui.

Para ler, estudar e compreender este livro, portanto, e para trabalhar com ele, não é necessário ter formação específica em lingüística. Mas é indispensável estar armado de duas coisas: motivação para aprender uma proposta nova de análise da estrutura do português; e, claro,

um pouco de tempo para ler, pensar, trabalhar exemplos etc. Desse modo, acredito que o leitor interessado conseguirá assimilar o essencial desta *Gramática* com seu próprio esforço.

Finalmente, pode valer a pena observar que, como proposta nova, este livro tem certo caráter polêmico. A todo momento pode tornar-se necessário confrontar a análise com sua concorrente tradicional, procurando mostrar a superioridade da primeira. Isso é inevitável, claro; mas não se espere encontrar aqui uma crítica completa e pormenorizada da análise tradicional, que me levaria além dos limites já um tanto amplos do tema do livro. O próprio leitor poderá, muitas vezes, fazer o confronto, exercício que me parece útil e instrutivo. Eu não gostaria de ver ninguém aceitar idéias (nem mesmo as minhas) sem saber por quê.

1.2. A LÍNGUA PADRÃO DO BRASIL

1.2.1. Variedades da língua

O primeiro problema a enfrentar quando se pretende elaborar uma gramática é o de definir que língua vai ser descrita. No caso do presente livro, isso pode parecer supérfluo: trata-se de uma gramática portuguesa e, portanto, a língua descrita é a portuguesa.

No entanto, o termo "língua portuguesa" é bastante ambíguo e se aplica a diversas variedades nitidamente diferenciáveis. Por exemplo, todos podemos distinguir claramente a fala de um português de Lisboa da de um brasileiro de Belo Horizonte; podemos igualmente distinguir a fala de um indivíduo criado na cidade, com segundo grau completo, da de uma pessoa criada no campo e analfabeta. E podemos também distinguir entre um texto escrito e a transcrição literal de um texto falado. Como exemplo desta última distinção, vejam-se os dois textos abaixo:

(1) A tarefa de lançar as bases da nova gramática é muito longa e complexa; devemos, portanto, deixá-la para a próxima semana.

(2) A nova gramática do português, ela vai ser muito difícil a gente escrever. Melhor a gente deixar ela pra semana que vem.

O exemplo (1) deve ter sido retirado de um texto escrito; já o exemplo (2) parece ser a transcrição de um trecho de fala espontânea. Não só as duas passagens estão organizadas de maneira diferente, mas há diferenças propriamente gramaticais entre elas: é muito raro, por exemplo, encontrar em textos escritos (principalmente textos técnicos) o pronome *ela* como objeto (*deixar ela*), ou a forma *pra* funcionando como contração de *p(a)ra* + *a* (*pra* [= *para a*] *semana que vem*).

Não é necessário entrar em detalhes acerca dessa diferença gramatical entre a língua escrita e a falada. Todos nós já tivemos contato com ela e sabemos que é uma das dificuldades que os escolares precisam enfrentar ao aprenderem a ler e escrever. Para efeitos deste livro, vou chamar a variedade ilustrada no texto (1) de "padrão", e a variedade do texto (2) de "coloquial".

Devo deixar bem claro que nem o padrão nem, principalmente, o coloquial são totalmente homogêneos. Existe uma imensa gama de variedades de língua, que vão desde as mais informais até as mais formais e estereotipadas. Além disso, há, principalmente no que se refere ao coloquial, certo grau de variação regional (a fala espontânea de um gaúcho difere da de um cearense) e social (um operário não fala da mesma maneira que um médico). Essa variação (regional, social e individual) é muito mais marcada no caso do coloquial do que no do padrão.

As diferentes variedades da língua são utilizadas em situações razoavelmente bem definidas. Assim, qualquer pessoa modifica sua maneira de falar conforme esteja discutindo no bar com os amigos, ou respondendo a uma entrevista para obter emprego. De modo geral, pode-se dizer que a variedade coloquial (ou melhor, o conjunto de variedades que chamamos "coloquiais") é utilizada na fala; já a variedade padrão é própria da escrita. Há exceções; por exemplo, o autor de um romance pode incluir muitos traços coloquiais nos diálogos, a fim de dar-lhes maior realismo; ou, inversamente, alguém pode tentar falar estritamente segundo o padrão ao fazer um discurso ou uma declaração em público. Mas essas situações são de certo modo marginais e, em geral, pode-se associar o padrão com a escrita, e o coloquial com a fala.

É importante assinalar que cada variedade tem seu conjunto de situações específicas e em geral não pode ser substituída por outras sem provocar estranheza ou mesmo destruir a possibilidade de comunicação. Imagine-se, por exemplo, a situação em que uma pessoa escreve um compêndio de economia e utiliza uma linguagem próxima da coloquial:

> A gente já falou aqui que pra agricultura desenvolver direito ela precisa mesmo ter uma porção de gente pra trabalhar no campo — quer dizer, aqui no Brasil, que não tem muita maquinaria agrícola, o trabalho tem que ser feito é na unha mesmo.

O texto vai chocar os leitores. Mesmo que o autor tenha uma mensagem inteligente a dar, é possível que ela fique perdida para muitas pessoas, simplesmente porque o compêndio não se adequa às convenções lingüísticas vigentes.

Por outro lado, a variedade padrão tampouco pode ser utilizada

em qualquer situação. Digamos que alguém chegue para seus amigos na mesa do bar e diga, textualmente, o seguinte:

> Amigos, sugiro que discutamos hoje a sensacional vitória do nosso clube na última partida, disputada no Mineirão, e que lhe atribuamos todo o mérito de que intimamente o sabemos credor.

A mensagem será interpretada como uma brincadeira. Mas, se a pessoa insistir em continuar falando dessa maneira (ou, pior, se resolver falar assim a vida toda), vai acabar sendo excluída da maioria de seus círculos de amizade. Ainda aqui, trata-se de simples convenção social; mas convenções sociais são coisas muito poderosas. Todos nós sabemos disso muito bem — tanto assim que estamos sempre ajustando nossa linguagem a cada situação.

O que se pode concluir daí é que cada variedade tem seus domínios próprios, onde é senhora quase absoluta. Não existe, simplesmente, uma variedade "certa". Cada situação de comunicação (ensaio científico, peça teatral, conversa de botequim, discurso de formatura, pedido de informações na rua etc.) impõe uma variedade própria, que é a "certa" naquela situação. É "errado" escrever um livro de economia em coloquial; mas é igualmente "errado" namorar ou conversar com os amigos utilizando o padrão. O leitor é convidado a fazer suas próprias observações; acredito que sairá convencido de que essa é a realidade.

1.2.2. O padrão brasileiro

Este livro é uma tentativa de descrição de uma variedade padrão da língua portuguesa. Por conseguinte, ocupa-se basicamente da língua escrita. Gostaria de acentuar que isso não significa que as variedades coloquiais não sejam dignas de estudo e que não seja urgente levar a efeito a tarefa de descrevê-las. Na verdade, tem-se feito muita pesquisa sobre o português coloquial, em suas variantes regionais, sociais, situacionais etc., mas talvez ainda seja cedo para se tentar uma síntese que possa ser chamada a "gramática" do português falado. Caminha-se nesse sentido, entretanto, e tal gramática terá, quando for feita, imenso interesse teórico e prático.

O presente livro não trata do português falado em nenhuma de suas variedades. Aqui vamos considerar apenas o português padrão escrito — a variedade, aliás, que é estudada pelas gramáticas tradicionais. Mesmo esse padrão admite variantes, e por isso vou deter-me na questão da definição da variedade a ser descrita, tentando defini-la com alguma objetividade. Retomo aqui as considerações do quinto capítulo de meu livro *Para uma nova gramática do português.*

Conforme apontei naquele volume, existe uma linguagem padrão utilizada em textos jornalísticos e técnicos (como revistas semanais, jornais, livros didáticos e científicos), linguagem essa que apresenta uma grande uniformidade gramatical, e mesmo estilística, em todo o Brasil. Assim, seria difícil distinguir lingüisticamente o editorial de um jornal de Curitiba do de um jornal de Cuiabá ou São Luís. Igualmente, a linguagem de um livro técnico ou didático publicado em Recife não se distingue da de um livro publicado em São Paulo ou Porto Alegre; os regionalismos não penetram em tais textos. Pode-se concluir que existe um português padrão altamente uniforme no país.

Esse padrão é encontrado em textos técnicos e jornalísticos em geral, mas nem sempre nos textos literários. O texto literário não apenas reflete as decisões pessoais do autor (que freqüentemente viola as regras do padrão a fim de obter efeitos estéticos), mas ainda pode conter traços do coloquial de uma ou outra região. Nesse sentido, pode-se identificar a região de origem de um autor (ou mesmo identificar o próprio autor) através de um exame de sua linguagem: Jorge Amado é baiano, e isso se reflete em sua linguagem, que nunca se confundiria com a de Érico Veríssimo, por exemplo.

Na elaboração da presente *Gramática*, optei por considerar o padrão geral, aquela variedade da língua que se manifesta de maneira uniforme nos textos técnicos e jornalísticos de todo o país. As razões foram expostas em meu livro acima citado, e aqui vou apenas mencioná-las: primeiro, tais textos oferecem uma uniformidade de estrutura que permite elaborar a descrição com maior coerência, pois ficam neutralizadas as inovações pessoais freqüentemente utilizadas em textos literários. Depois, parece-me que a própria linguagem literária costuma ser estudada em termos de desvios a partir de um padrão básico — o que pressupõe, naturalmente, um estudo prévio desse padrão.

As gramáticas usuais costumam trazer exemplificações tiradas predominantemente de obras literárias e, portanto, aparentemente partem de uma decisão diferente da minha quanto ao *corpus* da análise. Mas essa diferença é mais aparente do que real. Os exemplos literários são sempre selecionados de maneira a excluir idiossincrasias ou desvios — ou seja, acabam sendo "filtrados", e o filtro me parece ser baseado exatamente no padrão geral, não-literário, que me proponho estudar nesta *Gramática*. Por exemplo, embora Mário de Andrade seja uma das autoridades mencionadas na maioria das gramáticas, nem tudo o que ele escreveu pode ser incluído nelas, como, por exemplo,

> Me baste apenas lembrar... [Andrade, 1980, p. 5]

Esse posicionamento do pronome clítico *me* não é considerado padrão. Logo, prefere-se não ver que Mário de Andrade, um dos autores nos quais se baseia a análise, o utiliza com freqüência.

É, pois, como se as gramáticas se baseassem em um padrão literário expurgado de idiossincrasias pessoais, regionalismos e coloquialismos. O resultado, acredito, é muito semelhante ao padrão técnico-jornalístico que forma a base da presente descrição.

1.3. OBJETIVOS DO ESTUDO DA GRAMÁTICA

1.3.1. Para que estudar gramática?

Uma das motivações que me levaram a escrever esta *Gramática* foi a demanda por uma renovação no ensino de gramática portuguesa no primeiro e segundo graus, demanda manifestada com igual insistência por professores e alunos. Por isso, vou deter-me um pouco na discussão de uma pergunta não propriamente gramatical, mas antes de caráter pedagógico:

Para que se inclui o ensino de gramática no currículo de primeiro e segundo graus?

Mesmo que admitamos (contra as evidências) que nossos alunos aprendem um bocado de gramática durante sua vida escolar, que proveito tiram eles desse conhecimento?

Não quero dizer que cada disciplina do currículo deva necessariamente encontrar uma aplicação mais ou menos imediata na vida cotidiana ou profissional dos alunos; mais adiante desenvolverei esse ponto. Mas certamente é preciso justificar de alguma forma a presença de cada disciplina; a pergunta se reformula, então, como um pedido de justificativa para o caso dos estudos gramaticais. Vou a seguir dar minhas idéias sobre isso — idéias que têm condicionado todo o trabalho que venho realizando no campo da gramática e de seu ensino.

Concorda-se, geralmente, que o grande objetivo do ensino de língua portuguesa é levar os alunos a **ler** e **escrever** razoavelmente bem. Pergunto, então: será que o estudo da gramática pode ajudar na aquisição da leitura e da escrita?

Acredito que a resposta é negativa. Tenho encontrado pessoas que crêem na eficácia de um conhecimento de gramática para melhorar o desempenho naqueles dois campos fundamentais. Mas nunca podem apresentar evidência em favor de sua crença. Ninguém, que eu saiba, conseguiu até hoje levar um aluno fraco em leitura ou redação a melhorar sensivelmente seu desempenho apenas por meio de instrução gramatical. Muito ao contrário, toda a experiência parece mostrar que entre os

pré-requisitos essenciais para o estudo da gramática estão, primeiro, habilidade de leitura fluente e, depois, um domínio razoável da língua padrão (já que esta é o objeto das gramáticas disponíveis). Assim, para estudar gramática com proveito, é preciso saber ler bem — o que exclui a possibilidade de se utilizar a gramática como um dos caminhos para a leitura. Creio que o mesmo vale, *mutatis mutandis*, para a redação.

Somos forçados a concluir que o estudo de gramática não oferece um instrumento para atingir o grande objetivo da língua portuguesa no primeiro e segundo graus.

Será então totalmente inútil? Não necessariamente, é claro; e vou defender a idéia de que o ensino de gramática pode ter um papel importante na formação intelectual dos alunos. Para expor minha posição, começarei fazendo uma breve digressão sobre o conteúdo do ensino de primeiro e segundo graus.

1.3.2. Componentes do ensino

Consideremos os objetivos propriamente cognitivos, intelectuais, do ensino; ou seja, excluindo os automatismos, assim como a formação de atitudes. Vamos limitar-nos ao ensino enquanto transmissão (e criação) de conhecimento, sem negar que haja outros ingredientes importantes — aliás, como se verá, acabaremos discutindo também um aspecto da formação de atitudes. Ora, o ensino cognitivo pode ser analisado em três componentes, a saber:

(a) **Componente de aplicação imediata:** trata-se de conhecimentos que serão imediatamente úteis na vida profissional ou cotidiana dos alunos. Por exemplo, é preciso aprender as quatro operações para saber dar e receber o troco correto; é preciso ler fluentemente para adquirir as informações necessárias etc.

(b) **Componente "cultural"**: trata-se de conhecimentos sem nenhuma aplicação visível à vida prática, mas que nossa sociedade considera essenciais à formação do indivíduo. Por exemplo, aprendemos na escola que a Terra gira em torno do Sol, e não vice-versa. Essa informação não tem importância prática, mas todos concordariam que precisa ser fornecida. O que pensaríamos de uma escola que resolvesse voltar ao ensino do sistema geocêntrico de Ptolomeu? Aprendemos igualmente a história do Império Romano, da Idade Média e do Brasil; aprendemos as capitais dos países da Ásia e assim por diante. Muito poucas pessoas chegam a aplicar diretamente esses conhecimentos; mas ninguém pensaria seriamente em retirá-los dos programas. Esse tipo de informação, no que pese sua falta de sentido prático, é absolutamente indispensável à formação do cidadão moderno.

(c) **Componente de formação de habilidades:** finalmente, o ensino

tem também a responsabilidade de desenvolver nos alunos habilidades intelectuais de observação e raciocínio. É por isso que não nos limitamos, ao ensinar ciências, a comunicar os "fatos" tais como são vistos pelos cientistas modernos. Há sempre algum esforço no sentido de mostrar aos alunos a trilha que foi seguida para o descobrimento desses fatos: fazer observações, raciocinar e tirar conclusões juntamente com os alunos. O cidadão da nossa era não é primariamente uma pessoa que sabe muito, mas antes uma pessoa que sabe pensar por si só.

Algumas matérias concentram sua contribuição ao ensino em um ou dois desses componentes. O que devemos perguntar agora é se a gramática poderá encontrar aí o seu lugar.

1.3.3. Os objetivos do ensino de gramática

Vamos voltar aos três componentes, para ver se a gramática tem algo a oferecer no âmbito de cada um.

Comecemos com o componente de aplicação imediata. Se rejeitarmos a aplicação do ensino gramatical ao desenvolvimento da leitura e da escrita (que é o que sugeri acima), a importância da gramática para este componente será bastante reduzida. Não é nula, porém: digamos que alguém encontre, ao ler, uma palavra desconhecida e queira olhá-la no dicionário. Se a palavra for, digamos, *regurgitássemos*, surgirá um problema: *regurgitássemos* não consta em nenhum dicionário. O dicionário tem *regurgitar* e também *regurgitação*; qual dessas palavras é a relevante para se conhecer o sentido de *regurgitássemos*? É evidente que aqui precisamos aplicar informações de ordem gramatical: *regurgitássemos* é um "verbo", e os "verbos" aparecem no dicionário em sua forma de "infinitivo"; e o "infinitivo" encontrado no dicionário é *regurgitar*, e não *regurgitação*. Esse é um exemplo da aplicação de conhecimento gramatical a uma situação prática. Temos de concluir que o conhecimento de gramática tem aplicação imediata, embora talvez bastante modesta.

Passo agora ao segundo componente, ao qual chamei (por falta de termo melhor) "cultural". Aqui a argumentação é mais difícil, mesmo porque as pessoas em geral não estão devidamente conscientes da importância do componente cultural na formação dos estudantes. Mas creio que se pode defender que o estudo de gramática tem também algo a contribuir dentro do componente cultural.

Consideremos primeiro o quanto nossa sociedade é orientada para o conhecimento científico. O fato de alguém ser advogado ou locutor de rádio não o isenta da obrigação de ter alguma idéia do estado

da ciência. É isso que justifica a presença de estudos relativamente avançados de física, química, matemática e biologia no segundo grau. A imensa maioria dos alunos nunca aplicará esses conhecimentos diretamente em sua atividade profissional. Mas espera-se do cidadão de uma nação do século XX que saiba que os corpos se dilatam com o calor; que o sal é composto de sódio e cloro; que nossos corpos são compostos de células, cada uma das quais contendo núcleo e citoplasma etc. É uma multidão de informações cuja existência é pressuposta quando se encontra uma frase como "ele come comida sem sal porque não tolera o sódio", ou então "o terceiro mundo funciona como um satélite das nações industrializadas". Essa massa de informações científicas é parte importante do corpo de conhecimentos comuns às pessoas instruídas, e sua presença nos permite a comunicação no nível exigido pela sociedade moderna.

Desconfio que a importância desse componente é freqüentemente subestimada pelos planejadores de currículos e programas, a julgar pela ênfase habitual na "utilidade" dos conhecimentos transmitidos. Por outro lado, faltam estudos que nos permitam avaliar as dimensões e a importância do componente cultural no atual sistema de ensino.

Mas, voltando à gramática, é importante lembrar que ela é (ou deveria ser) uma aplicação da lingüística — uma ciência social, provavelmente a que mais desenvolvimento tem experimentado nos últimos tempos. Além disso, trata da linguagem, que é o mais básico dos fenômenos sociais, por permear todas as atividades de uma comunidade. Creio que a presença de uma formação lingüística em geral (e gramatical em particular) é defensável no ensino de primeiro e segundo graus exatamente pelas mesmas razões que justificam a preocupação de dar aos alunos alguma formação em química, biologia, história ou astronomia.

Além disso, o conhecimento da língua — e falo aqui do conhecimento explicitado de sua estrutura, não apenas do seu uso correto — é uma faceta importante do conhecimento da própria nação. Assim como a necessidade de preservar nossa cultura (e o amor-próprio nacional) nos leva a estudar a história, a geografia e o folclore do Brasil, não podemos deixar de estudar a língua que falamos (e a que escrevemos), em seus muitos aspectos: dialetológicos, históricos, sociais e também gramaticais. Sem estender-me mais sobre esse vasto e mal explorado assunto, chamo a atenção do leitor para esse importante aspecto do ensino.

Chegamos agora ao terceiro componente, a saber, a formação de habilidades intelectuais. É aqui que, acredito, os estudos gramaticais têm mais a oferecer; e é aqui que eles têm sido, na prática, mais nocivos. O assunto requer, portanto, nossa atenção cuidadosa.

Relembremos primeiro que as habilidades de raciocínio, de observação, de formulação e testagem de hipóteses — em uma palavra, de independência de pensamento — são um pré-requisito à formação de indivíduos capazes de aprender por si mesmos, criticar o que aprendem e criar conhecimento novo. É justamente neste setor que o estudo de gramática pode dar sua contribuição mais relevante; e é neste setor que nosso sistema educacional se tem mostrado particularmente falho: se há algo que nossos alunos em geral não desenvolvem durante sua vida escolar é exatamente a independência de pensamento. O estudante brasileiro (e, muitas vezes, também o professor) é tipicamente dependente, submisso à autoridade acadêmica, convencido de que a verdade se encontra, pronta e acabada, nos livros e na cabeça das sumidades. Daí, em parte, a perniciosa idéia de que educação é antes de tudo transmissão de conhecimento — quando deveria ser em primeiro lugar procura de conhecimento e desenvolvimento de habilidades.

Como é que o estudo gramatical poderia ajudar na formação dessas habilidades? Como se poderia dar a uma aula de gramática o caráter de um momento de pesquisa?

Nossa disciplina desfruta de uma situação favorável, se a comparamos com outras disciplinas científicas. Primeiro, os fenômenos estudados são relativamente fáceis de observar: não dependemos de laboratórios (como os químicos), nem de penosas viagens ao campo (como os geólogos) para apresentar aos alunos dados em primeira mão. Esses dados, basta elucidá-los dentro da sala de aula, pois estão programados no cérebro de cada falante da língua (não quero dizer que não existam sérios problemas metodológicos na obtenção de dados para a análise gramatical; mas é possível evitá-los em um estágio inicial do estudo). Em segundo lugar, existem teorias razoavelmente desenvolvidas de gramática, capazes de sugerir questões de profundidade variada: algumas, pelo menos, ao alcance da argumentação de um aluno de primeiro grau.

Se acrescentarmos que a linguagem, em todos os seus aspectos, permeia a vida social a todo momento, sendo um fenômeno altamente importante na vida das pessoas, veremos com clareza que a gramática oferece um campo privilegiado para o exercício das atividades de pesquisa. O estudo de gramática pode ser um instrumento para exercitar o raciocínio e a observação; pode dar a oportunidade de formular e testar hipóteses; e pode levar à descoberta de fatias dessa admirável e complexa estrutura que é uma língua natural. O aluno pode sentir que está participando desse ato de descoberta, através de sua contribuição à discussão, ao argumento, à procura de novos exemplos e contra-exemplos cruciais para a testagem de uma hipótese da-

da. Nesse sentido a gramática tem imensas potencialidades como instrumento de formação intelectual.

Concluo que a grande contribuição que o ensino gramatical encerra reside na possibilidade de ajudar o desenvolvimento das habilidades mencionadas; isto é, o ensino gramatical pode ser um dos meios pelos quais nossos alunos crescerão e se libertarão intelectualmente. Eu gostaria que este livro se tornasse, entre outras coisas, uma fonte de material de apoio para que o professor e os alunos realizassem essa potencialidade do estudo de gramática.

Esse objetivo só poderá ser atingido, evidentemente, se ficar perfeitamente claro e entendido que o estudo da gramática é parte da formação científica dos alunos; que trata da descrição, interpretação e compreensão de um aspecto do universo social que nos cerca; e, principalmente, que é um corpo de conhecimentos em constante revisão, sujeito a crítica, acréscimos e refutações. Uma gramática, enquanto descrição de uma língua, é na verdade um conjunto de hipóteses, mais ou menos bem fundamentadas. O mínimo que se pode fazer é conhecer a argumentação que está por trás da descrição proposta; sem isso, não se está estudando gramática. Em outras palavras, não se pode **estudar** gramática sem ao mesmo tempo **fazer** gramática.

Por isso, considero negativa a adoção oficial ou semi-oficial de uma nomenclatura (inseparável de uma análise igualmente oficial) para a gramática. Nomenclaturas como a NGB acabam sendo aceitas como obrigatórias — ou seja, a NGB, cuja motivação inicial foi provavelmente apenas a de evitar a proliferação de termos distintos para a mesma noção, tornou-se, na prática, uma **doutrina oficial**, fora da qual nenhum texto é aceitável para uso nas escolas. Isso não se admite em nenhuma outra disciplina e é igualmente absurdo em gramática. A única razão aceitável para a adoção de uma doutrina gramatical é o reconhecimento de que ela descreve os fatos com maior adequação do que suas concorrentes. Se isso acarretar, como é provável, alguma heterogeneidade no ensino gramatical do país, será um preço a pagar pela possibilidade de se utilizar a gramática como espaço para o exercício do pensamento. Um preço, acredito eu, muito baixo, em vista da qualidade da mercadoria.

O grande problema da adoção da NGB e de uniformizações oficiais em geral não é a sua inadequação enquanto descrição da estrutura da língua, mas sua aceitação passiva como doutrina oficial, acima de questionamentos e reformulações. A mudança a empreender com urgência não está na gramática, mas na atitude geral frente à gramática. Substituir a NGB por uma descrição mais correta e atualizada é importante, mas não é o mais importante; conseqüentemente, a proposta desta *Gramática* só se compreende se for enca-

rada de maneira não-dogmática. Só assim este trabalho poderá atingir seus objetivos dentro do ensino.

1.3.4. Gramática normativa

Antes de encerrar esta seção, vou fazer uma advertência a respeito do ensino normativo da língua. Defendo a idéia de que o ensino normativo não é um mal em si, mas que tem sido aplicado (como tanta coisa no nosso campo) de maneira prejudicial aos alunos.

Já vimos que o estudo de gramática não é um caminho conveniente para desenvolver o desempenho na leitura e na escrita (muito menos na fala). Outra coisa que a gramática não deveria ser é um instrumento de ensino normativo. Aqui, sua ação tem sido mais que inútil, tem sido desastrosa.

O grande perigo é transformar a gramática — uma disciplina já em si um tanto difícil — em uma doutrina absolutista, dirigida mais ou menos exclusivamente à "correção" de pretensas impropriedades lingüísticas dos alunos. A cada passo, o aluno que procura escrever encontra essa arma apontada contra sua cabeça: "Não é assim que se escreve (ou se fala)", "Isso não é português" e assim por diante. Daí só pode surgir aquele complexo de inferioridade lingüístico tão comum entre nós: ninguém sabe português — exceto, talvez, alguns poucos privilegiados, como os que se especializam em publicar livros com listas de centenas ou milhares de "erros de português".

Acaba-se desencorajando nos alunos qualquer iniciativa no campo da linguagem, em especial no da escrita. Todos conhecemos a figura do adulto bastante bem-educado, com formação profissional e intelectual adequada, e que teria algo a dizer, mas que não se arrisca a escrever uma página para publicação, ou mesmo um simples relatório, justamente porque "não sabe português". Se o professor de português tem diante de si um desafio, é o de ir contra essa tradição repressiva, que deforma a imagem de nossa disciplina aos olhos de alunos e ex-alunos. Essa imagem de "polícia da linguagem" se manifesta a todo momento, por exemplo quando, ao saberem que trabalhamos com pesquisa em língua portuguesa, sempre acabam nos perguntando se esta ou aquela construção é a "certa". Isso pode nos dar alguma pequena parcela de poder, é certo; mas não contribui para que nos estimem, nem para que nos respeitem enquanto trabalhadores intelectuais.

Não quero dizer com isso que o ensino normativo deva ser suprimido. É preciso, apenas, colocá-lo em termos mais realistas. Não se trata de confrontar um "português certo" e um "português errado", definidos ambos em termos absolutos, independentemente do contexto si-

tuacional ou social. Trata-se, antes, de defender a idéia de que a cada situação corresponde uma variedade distinta da língua (ver a seção 1.2.1. acima).

Concordo, portanto, que é necessário ensinar o português padrão; mas esse ensino (o "ensino normativo" da língua) deve ser atacado com muita cautela e com toda a diplomacia. Como qualquer material potencialmente explosivo, deve ser manejado com cuidado. Acredito que, com o desenvolvimento da leitura fluente e do hábito da leitura, a maior parte do problema acaba desaparecendo por si só. O que restar poderá ser objeto de ensino, mas sempre deixando bem explícito o verdadeiro papel do padrão escrito frente às variedades coloquiais. Não levar isso em conta resultará, fatalmente, em agravar os já sérios problemas que infestam nossa disciplina, o ensino em geral e, em última análise, a vida de nossos alunos.

2 Princípios do estudo de gramática

2.1. METODOLOGIA

Antes de atacar o estudo da gramática portuguesa, é importante ter alguma noção dos princípios em que se baseia a descrição das línguas. Esse preliminar é particularmente importante porque nossa tradição de estudo gramatical freqüentemente despreza a fundamentação teórica e metodológica — e, como resultado, apresenta uma descrição cheia de inconsistências e às vezes sem grande conexão com a língua real. Assim, embora nosso objetivo nesta *Gramática* seja estudar a estrutura da língua portuguesa, é indispensável dedicar alguma atenção a certos pontos teóricos. Neste capítulo vou apenas esboçar, muito sumariamente, os pontos mais importantes; mas o estudioso de gramática deverá manter um constante interesse em questões de teoria e metodologia. A qualidade da gramática depende crucialmente de uma atenção permanente a tais questões. Dizendo a mesma coisa de outro modo (e correndo o risco de ser repetitivo), direi ainda que **o estudo da gramática de uma língua não pode dispensar o estudo da teoria e da metodologia lingüísticas.**

Além dos pontos de metodologia abordados neste capítulo, vários outros são tratados à medida que se tornam necessários; por exemplo, no capítulo 12, discuto os princípios em que se baseia a classificação das palavras.

2.1.1. Os dados

Conforme vimos no capítulo 1, esta *Gramática* é uma descrição do português padrão brasileiro, a modalidade de língua utilizada no Brasil na maior parte dos textos escritos. Boa parte da população tem algum domínio dessa modalidade do português: em geral, as pessoas funcionalmente alfabetizadas conseguem compreender adequadamente o padrão, embora nem sempre o utilizem corretamente na escrita. Mas é

importante notar que ninguém no Brasil tem o português padrão como sua língua nativa. Todos nós, mesmo os mais escolarizados, aprendemos em primeiro lugar a modalidade coloquial, que difere em pontos importantes do padrão; e todos nós utilizamos o coloquial na maioria das situações de fala. Essa situação, como veremos, condiciona em parte o problema da obtenção dos dados para a análise.

O método tradicional em sintaxe para a obtenção de dados depende principalmente do julgamento dos falantes sobre frases isoladas. Às vezes, pode-se também recorrer ao exame de textos, mas esse recurso é impraticável para o todo da análise sintática, porque implicaria na manipulação de quantidades imensas de material. Assim, tipicamente, o lingüista constrói exemplos e utiliza seu próprio julgamento quando este é claro; caso contrário, pode também testá-lo mais ou menos sistematicamente com outros falantes. E nada impede que recorra ao exame de textos para a verificação de pontos específicos.

Essa metodologia tem seus inconvenientes, mas até agora não se conhece uma melhor. Por exemplo, freqüentemente o lingüista se depara com julgamentos vacilantes: alguns falantes aceitam uma construção como bem formada; outros não têm muita certeza; ainda outros a rejeitam. Essa variação é, na verdade, mais um fato lingüístico, mas que não pode ser levado em conta, a não ser em escala limitadíssima, em uma descrição como esta. Como resultado, a imagem da língua aqui delineada é inevitavelmente aproximativa. Seria desejável elaborar uma metodologia para elucidar dados, interpretar a variação de julgamento e incorporá-la na descrição. Esse é um trabalho que está sendo atacado por alguns lingüistas (notadamente na área de sociolingüística); mas por ora uma gramática como a presente não se pode beneficiar dessa metodologia. Fiquemos, pelo menos, conscientes das limitações da descrição apresentada.

O estudioso de gramática tradicional não costuma sentir essa incerteza. As gramáticas transmitem, implicitamente, a idéia de que é bem nítida a diferença entre as construções do padrão (chamadas "corretas") e as que não pertencem ao padrão, seja por não existirem na língua, seja por pertencerem ao coloquial (chamadas "erradas", "incultas" etc.). Mas isso é uma ilusão, mantida através de certos recursos que são rejeitados neste trabalho.

Um desses recursos é simplesmente fechar os olhos às construções que, embora ocorrendo nas obras dos "autores credenciados", não se coadunam com a descrição. Assim, é preciso "expurgar" boa parte da obra de escritores como Mário de Andrade, Jorge Amado, Luís Fernando Veríssimo, Luís Vilela — na

verdade, talvez a maioria dos autores mais representativos da época atual.

Outro recurso é simplesmente decidir arbitrariamente que algo é aceitável ou não, sem verificação isenta. Assim, por exemplo, se perguntado sobre a "correção" da frase

(1) ? Os únicos três sobrinhos de Cláudio moram juntos.

um gramático ou um professor de tendência tradicional poderá "aprová-la" ou não, sem mencionar o fato (pois é um fato) de que muitos falantes bem instruídos no padrão hesitam em seu julgamento. Naturalmente, essa atitude arbitrária é facilitada pela crença bastante espalhada de que uma construção é "correta" (pertence ao padrão) **porque** é aprovada pelos gramáticos, o que a meu ver coloca definitivamente o carro adiante dos bois (ver a respeito a seção 1.3.4.).

Vacilações como a que encontramos no julgamento de (1) são, por assim dizer, fatos da vida: ossos do ofício gramatical, incômodos mas reais. Não aprenderemos nada escondendo a realidade dos outros e muito menos de nós mesmos. Melhor, creio, será reconhecer que a língua é um instrumento de enorme complexidade, utilizado a todo momento por milhões de pessoas que diferem quanto a grau de conhecimento, experiência lingüística prévia, crenças, gostos e preferências. É de esperar que haja variação; e, se uma descrição como a proposta nesta *Gramática* não faz jus a toda a variação observada, isso se deve apenas a limitações de tempo, espaço e nível de conhecimento do objeto. Em outras palavras, uma gramática só pode apresentar um compromisso aceitável entre o retrato fiel da realidade (o que a complicaria além do razoável) e a simplificação excessiva (o que lhe tiraria a relevância). É esse difícil compromisso que será tentado no presente trabalho.

A alternativa, é preciso repetir, não é uma gramática completa, mas apenas a **ilusão** de uma gramática completa. Conforme o leitor verificará, em muitos pontos a análise apresentada neste livro deixa questões em aberto, seja porque ainda não foram devidamente pesquisadas, seja em virtude de problemas inerentes aos próprios dados. Por exemplo, veremos no capítulo 4 que a classificação exata de certos elementos internos do sintagma nominal — por exemplo, a classificação da palavra *cada* — é impossível de averiguar com certeza. A razão é que esses elementos são de ocorrência tão restrita que não se podem construir os testes que determinariam com precisão seu lugar na taxonomia dos itens que compõem o sintagma nominal.

Esse problema não pode ser evitado — pode apenas ser ocultado, digamos por meio da colocação de *cada*, arbitrariamente, em uma ou outra classe dentre as bem definidas. Mas o que se ganhará com isso? Nosso objetivo é atingir alguma com-

preensão da estrutura da língua — e isso inclui, claro, a identificação daqueles pontos nos quais, por razões diversas, a descrição não se pode fazer tão detalhadamente quanto desejaríamos. Não vejo razão para nos angustiarmos diante dessa situação; todas as ciências enfrentam o mesmo tipo de problema, pois o acesso aos dados relevantes é sempre limitado. Hoje não se tem certeza do número de satélites de Plutão, e não se conhece bem a geografia da Antártida; isso não desmerece a astronomia ou a geografia, nem os cientistas que se dedicam a essas disciplinas. Por que deveria ser diferente com a gramática?

2.1.2. Forma e significado

As unidades lingüísticas apresentam, como se sabe, dois aspectos fundamentais: a forma (ou "significante") e o significado. A forma e o significado não são propriamente partes ou componentes das unidades lingüísticas, mas antes aspectos evidenciados pelo ponto de vista adotado ao se estudar a língua.

Assim, uma unidade como a palavra *reloginhos* pode ser estudada do ponto de vista **formal**, e nesse caso levar-se-á em conta: sua pronúncia; sua composição morfológica (radical *relog-*; sufixo *-inho*; sufixo *-s*); e seu comportamento sintático (admite a anteposição do artigo *os*, mas não do artigo *as*; pode constituir o núcleo de um sintagma nominal; não concorda em pessoa etc.). Por outro lado, a mesma palavra *reloginhos* pode ser estudada do ponto de vista **semântico** (isto é, do significado); nesse caso teremos de levar em conta seu significado básico ("objeto para indicar as horas"); o fato de denotar objeto relativamente pequeno; o fato de indicar mais de um desses objetos etc.

Os dois aspectos, o formal e o semântico, estão presentes na palavra *reloginhos*, mas precisam ser separados na descrição. Essa separação é fundamental quando se estuda a gramática, porque a relação que existe entre as formas gramaticais e o significado que elas veiculam é extremamente complexa e indireta. Na verdade, a explicitação dessa relação é um dos objetivos primordiais da análise lingüística — e por isso mesmo é essencial descrever os dois aspectos separadamente, para depois colocá-los em confronto.

Em um livro anterior (Perini, 1989, p. 19-27), discuti esse assunto em algum detalhe. Aqui vou apenas dar alguns exemplos, remetendo o leitor àquele livro para uma exposição mais completa. Relembro apenas que a separação metodológica entre forma e significado não é nenhuma inovação; trata-se de um princípio aceito há muitas décadas, por grande parte dos lingüistas.

Digamos que nos interessa estudar a função denominada "sujeito". Na frase abaixo

(2) Joanita plantou jerimum no jardim.

diz-se que *Joanita* é o sujeito. Conforme veremos nos capítulos seguintes, o sujeito é o elemento da oração que está em relação de concordância com o verbo. Ou seja, a palavra *plantou* está nessa forma particular (com a terminação *-ou*) porque *Joanita* é um termo de terceira pessoa, singular. Se em vez de *Joanita* colocarmos *nós*, o verbo terá de mudar de forma:

(3) Nós plantamos jerimum no jardim.

Essa é uma observação que diz respeito à **forma** da frase. Note-se que não se disse nada a respeito do que é que *Joanita* quer dizer na frase; só notamos uma relação formal dessa palavra com a palavra *plantou*.

Outra observação formal que podemos fazer é que *Joanita* está no início da frase; ou que é um termo formado de uma única palavra. Todas essas afirmações expressam aspectos formais da frase (2) e da maneira como *Joanita* se insere nessa frase.

Agora vejamos algumas afirmações de caráter semântico. Primeiro, podemos observar que *Joanita*, ao que tudo indica, se refere a uma pessoa do sexo feminino. Se tivéssemos *Ricardo*, seria uma pessoa do sexo masculino; e, se tivéssemos *minhas tias*, seriam várias pessoas. Outra observação é que *plantou* exprime uma ação, e *Joanita* exprime a pessoa que praticou essa ação (o **agente** da ação).

Como se vê, é possível (e indispensável) descrever os aspectos formais e os semânticos separadamente. Por outro lado, é igualmente importante colocá-los em correlação. No caso, já verificamos que *Joanita* é sujeito da frase (2); podemos dizer então que nessa frase o sujeito exprime o agente da ação expressa pelo verbo.

Essa afirmação só faz sentido porque já dispomos de uma caracterização **formal** de "sujeito" (elemento com o qual o verbo concorda) e uma caracterização **semântica** de "agente" (entidade que pratica uma ação). Por isso insisto na necessidade de estudar separadamente os dois aspectos. Suponhamos, por exemplo, que tivéssemos definido "sujeito" utilizando a noção de "entidade que pratica a ação"; nesse caso, não poderíamos apresentar como um fato interessante da língua a coincidência, em (2), do sujeito e do agente — porque isso não seria uma informação real sobre a língua, mas uma decorrência automática da nossa definição de "sujeito". Ou seja, se "entidade que pratica a ação" é parte da definição de "sujeito", então **necessariamente** todo sujeito denota a entidade que pratica a ação (de outro modo não seria sujeito, segundo a definição dada).

Esse problema se generaliza para todas as noções utilizadas na análise da língua: funções sintáticas, classes e processos. Só podemos descrever a relação entre forma e significado se descrevemos primeiramente, e separadamente, a forma e o significado.

Se a relação entre a forma e o significado fosse sempre simples como a que acabamos de ver na frase (2) — isto é, o sujeito exprime o agente —, talvez fosse admissível esquecer um pouco a necessidade de separar os dois aspectos quando da análise. Mas a relação é das mais complexas; por exemplo, não se pode concluir do exemplo (2) que o sujeito sempre exprime o agente de uma ação. Primeiro, há muitos verbos que não exprimem ação (*morrer, amar, ser, parecer* etc.); depois, há verbos que exprimem ação, mas o agente não é expresso pelo sujeito, como em

(4) Eu apanhei de Joanita.

Eu é o sujeito (o verbo concorda com ele); a frase exprime uma ação; mas essa ação foi evidentemente praticada por Joanita, e a palavra *Joanita* não é o sujeito de (4).

Isso só serve como ilustração; ver o capítulo 10, onde são discutidos numerosos exemplos de relação entre certas funções sintáticas e aspectos do significado. Acredito que o leitor ficará convencido, se não da adequação das soluções ali propostas, pelo menos da complexidade da relação entre forma e significado.

Entende-se, portanto, a descrição de uma língua como composta essencialmente de três componentes: uma descrição formal; uma descrição semântica; e, finalmente, um sistema que relaciona o plano semântico com o plano formal. A descrição formal corresponde ao que chamamos **fonologia**, **morfologia** e **sintaxe**; e o sistema que correlaciona os dois planos é composto das **regras semânticas** (ou, mais exatamente, **regras de interpretação semântica**), estudadas nos capítulos 9 a 11 desta *Gramática*.

A fonologia, a morfologia e a sintaxe são igualmente compostas de regras. Pode-se conceber a diferença entre essas regras e as regras semânticas da seguinte maneira: as regras fonológicas, morfológicas e sintáticas definem quais são as construções possíveis na língua — ou seja, entre outras coisas explicam por que construções como * *port*, * *corrimos* e * *eu chegamos agora* são inaceitáveis[1]: a primeira viola uma regra fonológica (que proíbe palavras terminadas em *t*); a segunda viola uma regra morfológica (o verbo *correr* está conjugado incorretamente); e a terceira viola uma regra sintática (concordância verbal).

Já as regras semânticas fornecem a relação entre as construções

[1] O asterisco (*) é utilizado para marcar palavras ou construções inaceitáveis, isto é, rejeitadas como mal formadas pelos falantes da língua.

da língua e seus significados — sua relação com o mundo extralingüístico. Assim, uma regra semântica estipula que, em *Pedro beliscou o touro*, o agente é Pedro e o paciente é o touro; em *o touro beliscou Pedro*, o papel de agente é atribuído ao touro, e o de paciente a Pedro.

Em geral, pode-se dizer que os traços formais da língua sofrem interpretação, ou seja, a cada traço formal corresponde algum traço de significado. Mas há traços formais que não têm nenhuma interpretação. Por exemplo, certos verbos têm complemento sem preposição, como em

(5) Meu filho detesta alface.

Outros exigem uma preposição:

(6) Meu filho gosta *de* alface.

Essa exigência ou não de preposição antes do complemento verbal é um traço puramente formal; nenhuma diferença semântica acompanha a diferença entre a presença ou ausência da preposição. Este é, portanto, um fenômeno que só pode ser estudado no plano formal.

Por outro lado, há igualmente diferenças semânticas que não correspondem a nenhuma diferença formal. Para tomar um exemplo bem claro, vejamos as frases

(7) Esta é a mulher mais bonita *de Belo Horizonte.*

(8) Esta é a poesia mais bonita *de Mário Quintana.*

(9) Esta é a gravata mais bonita *de Mário Quintana.*

Os elementos grifados nessas três frases têm exatamente as mesmas propriedades formais: a mesma estrutura interna e o mesmo tipo de relacionamento com o restante da construção. No entanto, as diferenças semânticas são evidentes: em (7), *de Belo Horizonte* exprime um lugar; em (8), *de Mário Quintana* exprime o autor da poesia; e em (9) o mesmo termo exprime o possuidor da gravata. Nos casos de (7), (8) e (9), não há nenhuma diferença formal a descrever: as três frases são, formalmente falando, idênticas (à parte, claro, a diferença de itens léxicos). Mas as diferenças semânticas são claras.

Como é evidente, casos como o de (5)-(6) não têm interesse para a semântica, mas precisam ser estudados na sintaxe. Ao contrário, o caso de (7)-(9) não tem interesse sintático, mas precisa ser levado em conta na semântica.

Mas, como disse, esses não são casos típicos, nem particularmente interessantes. Os exemplos mais reveladores (e também os mais difíceis de estudar) são aqueles em que uma diferença formal corresponde a uma diferença semântica. Vários desses casos serão estudados em bastante detalhe nos capítulos da terceira parte deste livro.

Sumariando esta seção, relembro que a descrição de uma língua (a gramática) inclui uma descrição de seus aspectos formais, mais um sistema de regras que relaciona esses aspectos formais aos significados que eles eventualmente veiculem. Os aspectos formais serão estudados principalmente nos capítulos 3 a 8, e as regras semânticas nos capítulos 9 a 11. Há também, evidentemente, o estudo das categorias de significado em si, mas esse estudo está fora do escopo desta *Gramática*.

Finalmente, enfatizo mais uma vez a necessidade de proceder ao estudo dos dois planos separadamente: ao estabelecer uma categoria formal, não é lícito lançar mão de noções semânticas; e, ao estabelecer categorias semânticas, não se pode utilizar noções formais. Em um segundo momento, as categorias assim independentemente estabelecidas são colocadas em confronto através das regras de interpretação semântica.

2.1.3. "Fatos" em gramática

2.1.3.1. O problema; "fatos sintáticos"

Um problema que se apresenta de imediato ao tentarmos estudar uma língua é o de selecionar os fatos que serão relevantes para a análise e os que deverão ser deixados de lado.

Talvez isso surpreenda algumas pessoas: afinal, não se deve estudar **tudo**? Por que deixar certos fatos de fora? A resposta é que nem todos os fatos observáveis nos enunciados de uma língua interessam à lingüística; e, além do mais, mesmo dentre aqueles que interessam à lingüística, nem todos são descritos na sintaxe (ou na morfologia, ou na semântica etc.). Por exemplo, tomemos o enunciado seguinte:

(10) A Silvinha não se pinta.

Agora examinemos alguns fatos facilmente observáveis nesse enunciado:

(a) tem dezenove letras;

(b) do ponto de vista métrico, forma um verso de redondilha maior;

(c) tem cinco palavras;

(d) *pinta* concorda em número e pessoa com *Silvinha*;

(e) *se* está entre *não* e *pinta*;

(f) a primeira letra de *Silvinha* tem a forma de uma curva reversa, arredondada.

Desses seis fatos, alguns são sempre levados em conta nas gramáticas: o fato (d) é tratado no capítulo da "concordância verbal"; o fato (e) no da "colocação dos pronomes átonos". Os lingüistas consideram relevante para a descrição o fato de que *Silvinha* de certo modo "governa" a forma de *pinta*, de tal modo que não se pode dizer em português * *a Silvinha não se pintamos*. Da mesma for-

ma, a posição da palavra *se* é governada por certas restrições, e não se pode dizer * *se a Silvinha não pinta*, nem * *a se Silvinha não pinta.* O fato (c) também pode ser tratado lingüisticamente, embora não exatamente na gramática; em geral é parte da lexicologia. Mas a gramática depende dessa separação lexicológica do enunciado (10) em cinco palavras, porque caso contrário poderíamos tentar estudar o comportamento gramatical da seqüência *vinha nã*, que afinal de contas está presente em (10). Essa seqüência não é estudada lingüisticamente porque há boas razões para segmentar (10) de maneira diferente em unidades gramaticalmente significativas (chamadas "palavras").

Agora considere-se o fato (b): pode ser relevante se (10) for parte de um poema, para descrever a métrica do mesmo. Mas isso já não se coloca dentro do campo de interesse da gramática; está no campo da poética (teoria da literatura, semiótica). Por isso, o lingüista simplesmente não o leva em conta. Também o fato (f) fica fora dos estudos gramàticais e irá interessar talvez a quem se ocupe da história do alfabeto ou da tipografia.

Como se vê, não seria possível (nem em princípio) elaborar uma gramática que incluísse a descrição de **todos** os fatos observáveis nos enunciados de uma língua. A gramática (como as demais disciplinas científicas) começa a trabalhar recortando no mundo dos fatos uma área própria de interesse.

Por isso faz sentido perguntar que tipos de fatos são relevantes para a análise gramatical. Aqui vou apenas relacionar uma série de fatos que são tradicionalmente levados em conta na descrição gramatical, e que serão igualmente levados em conta na presente proposta. Uma discussão mais detalhada do assunto se encontra em Perini, 1989, p. 35-56.

Vamos concentrar-nos nos fatos que interessam à sintaxe (tal como se define tradicionalmente o campo da sintaxe). Não se trata propriamente de "fatos sintáticos", porque os fatos não vêm rotulados; colocar um fato na sintaxe já é uma decisão derivada da teoria. Mas há um conjunto de fatos que se colocam na sintaxe por consenso mais ou menos geral dos lingüistas, e esses se chamam, na gíria do ofício, "fatos sintáticos".

Os fatos sintáticos mais importantes pertencem aos seguintes tipos:

(a) posição linear na seqüência;

(b) agrupamento em constituintes;

(c) manifestações da relação de regência;

(d) correspondência;

(e) retomada anafórica.

A seguir vou conceituar e exemplificar cada um dos tipos.

2.1.3.2. *Posição linear*

A **posição linear** é simplesmente a posição que uma unidade ocupa em relação às outras unidades do enunciado. Assim, podemos dizer que a posição de *Silvinha* em (10) é em segundo lugar, ou logo antes de *não.*

Muitas vezes é necessário formular propriedades sintáticas de um item (por exemplo, uma palavra) em termos da posição linear: vimos que a palavra *se* em (10) só pode ocorrer entre *não* e *pinta,* e essa limitação é estabelecida pela gramática. Em outros casos, a posição pode variar, dentro de certos limites também estabelecidos pela gramática. Um exemplo é a posição de *todos* nas frases seguintes:

(11) a. Todos os motoristas entraram em greve.
b. Os motoristas todos entraram em greve.
c. Os motoristas entraram todos em greve.
d.* Os todos motoristas entraram em greve.
e.* Os motoristas entraram em todos greve.

2.1.3.3. *Constituintes*

Entendemos por **constituintes** certos grupos de unidades que fazem parte de seqüências maiores, mas que mostram certo grau de coesão entre eles. Na frase

(12) A casa de Lulu é azul e branca.

os falantes "sentem" que *a casa de Lulu* forma uma unidade, o que não se verifica com *Lulu é azul.* Dizemos então que *a casa de Lulu* é um **constituinte** e que *Lulu é azul* [na frase (12)] não é um constituinte.

A idéia é que as frases são formadas de constituintes, muitas vezes uns dentro dos outros. Assim, a frase (12) poderia ser analisada como contendo, entre outros, os constituintes seguintes:

a casa de Lulu é azul e branca
a casa de Lulu
casa de Lulu
azul e branca
é azul e branca etc.

Note-se que certos constituintes estão dentro de outros: o constituinte *a casa de Lulu* está dentro do constituinte *a casa de Lulu é azul e branca,* e o constituinte *azul e branca* está dentro do constituinte *é azul e branca,* que por sua vez está dentro de *a casa de Lulu é azul e branca.* Note-se que a frase completa é igualmente um constituinte.

Essa estruturação é freqüentemente mostrada através de um diagrama em árvore, assim:

(13)

É importante ter uma boa noção da estruturação das frases em

constituintes, porque toda a análise se baseia nela. Por exemplo, os constituintes costumam receber uma "função" na análise tradicional: *a casa de Lulu* é "sujeito", e *azul e branca* é "predicativo do sujeito". Já a seqüência *Lulu é azul* não recebe função nenhuma, pois não é um constituinte.

Os falantes têm muitas vezes intuições bem definidas sobre os constituintes; em outros casos as intuições não são claras, e é preciso lançar mão de outros recursos, como tentar observar o comportamento sintático de uma seqüência em outras frases. Esses problemas serão enfrentados à medida que ocorrerem; por ora, é importante ficar conhecendo a noção geral de "constituinte" e sua função na estruturação das frases.

2.1.3.4. A relação de regência

É tradicional considerar que em certos casos um constituinte "rege" outro, no sentido de que determina em parte a forma desse outro. Assim, nos casos de concordância verbal, por exemplo, diz-se que o verbo "concorda" com o sujeito: ou seja, o sujeito regeria o verbo.

Em português, o fenômeno da regência se manifesta principalmente sob três formas: concordância (verbal ou nominal), transitividade e ocorrência de pronomes oblíquos.

Os casos de concordância são bem conhecidos e não necessitam de comentário; trata-se simplesmente da exigência de que, quando dois termos compõem um constituinte maior, um deles (o "regido") assuma determinada forma para harmonizar-se com certos traços do outro (o "regente"). No exemplo

(14) Uma casa amarela

casa é marcado como "feminino" e está no singular. Como *casa* rege *uma* e *amarela*, estas últimas precisam ocorrer também no feminino singular (*uma*, *amarela*; e não, por exemplo, *um*, *amarelo*, nem *umas*, *amarelas*). O caso da concordância verbal (verbo e sujeito) é semelhante. O que nos interessa observar aqui é como um constituinte determina em parte a forma do outro.

Outro aspecto do que se considera o mesmo fenômeno é a exigência que o verbo pode fazer quanto à forma de seu complemento. Consideremos os verbos *ver*, *pensar* e *gostar*. Todos eles podem ter complementos, mas com *ver* esse complemento é um sintagma nominal, como em

(15) Amélia viu um passarinho.

Com *pensar*, o complemento precisa vir precedido da preposição *em*:

(16) Amélia pensou no passarinho.

E com *gostar* o complemento precisa vir com a preposição *de*:

(17) Amélia gostou do passarinho.

Novamente, como se vê, um constituinte determina a forma de outro.

Finalmente, temos o caso do aparecimento de pronomes oblíquos. Esses pronomes (*me, mim, o, nos, os*) ocorrem em certas circunstâncias que são, em última análise, regidas por um verbo ou uma preposição. Tomando como exemplo os pronomes de primeira pessoa: com o verbo *ver* o complemento (objeto direto) precisa ser *me*:

(18) Amélia me viu na feira.

Mas se o pronome ocorrer em um constituinte com a preposição *de* deverá assumir a forma *mim*:

(19) Amélia desistiu de mim.

E se a preposição for *com* o pronome será *-migo* (tradicionalmente escrito junto com a preposição como uma só palavra):

(20) Amélia se zangou comigo.

Podemos então dizer que o verbo e a preposição "regem" o pronome.

Esses são os casos em que a relação de regência é diretamente observável em português. Também se admite relação de regência em certos casos em que não há manifestação morfossintática explícita; não trataremos de tais casos aqui, pois só nos interessa no momento estabelecer a noção de "regência" com clareza.

2.1.3.5. Correspondência

Na maioria das análises considera-se que a frase

(21) O gato comeu o rato.

tem um relacionamento especial com

(22) O rato, o gato comeu.

mas não se relaciona da mesma forma com

(23) O rato comeu o gato.

Uma razão para isso, evidentemente, é a semelhança de significado entre (21) e (22), que são, até certo ponto, sinônimas. No entanto, pode-se também definir uma afinidade **formal** (sintática) entre (21) e (22), que as separa de (23). Diremos aqui que as frases (21) e (22) são **correspondentes**, ou que se **correspondem**. Mas nem (21) nem (22) corresponde a (23).

A relação de correspondência é extremamente importante na sintaxe atual, embora nem sempre seja explicitamente definida. Considera-se que a existência de uma relação de correspondência entre duas frases [como entre (21) e (22)] é um "dado bruto" da análise — ou seja, mais um tipo de "fato sintático" que pode ser utilizado, por exemplo, em definições. Assim, é importante definir a relação de correspondência nesta *Gramática*. Antes disso, porém, darei alguns exemplos de grupos de estruturas correspondentes. Um grupo de estruturas relacionadas pela

correspondência pode ser chamado um **grupo de correspondência**.

Entre os grupos de correspondência existentes em português podemos mencionar:

Topicalização

(21) O gato comeu o rato.
(22) O rato, o gato comeu.

Como se vê, as frases se diferenciam porque uma delas tem um dos termos colocado no início da oração, freqüentemente separado por vírgula. Esse elemento se diz **topicalizado**.

O português permite a topicalização de muitos termos da oração:

(24) a. O gato comeu o rato rapidamente.
b. Rapidamente, o gato comeu o rato.

(25) a. Os alunos trouxeram a maçã para a professora.
b. Para a professora, os alunos trouxeram a maçã.
c. A maçã, os alunos trouxeram para a professora.

Como se pode observar bastante claramente no grupo (25), frases correspondentes não são necessariamente sinônimas. As frases de (25) descrevem a mesma situação, mas o fazem tomando elementos diferentes como ponto de partida; em outras palavras, as frases de (25) "afirmam alguma coisa" a respeito de entidades diversas. O elemento que exprime essa entidade é o elemento topicalizado — assim chamado por ser o **tópico** (o assunto principal) do enunciado.

Clivagem

Outro recurso que o português possui para colocar termos em realce é a **clivagem**, que também dá origem a grupos de correspondência:

(26) a. O gato comeu o rato.
b. Foi o gato que comeu o rato.
c. Foi o rato que o gato comeu.

Como se vê, as estruturas clivadas se formam com o auxílio do verbo *ser* mais *que* (às vezes *quem*), além da anteposição do termo clivado.

"Advérbios" transpostos

Sabemos que certos elementos (tradicionalmente chamados "advérbios") podem ocorrer em diversas posições na oração. Naturalmente, resultam daí também grupos de correspondência:

(27) a. Aparentemente, Aldo gostaria de mudar de emprego.
b. Aldo, aparentemente, gostaria de mudar de emprego.
c. Aldo gostaria, aparentemente, de mudar de emprego.
d. Aldo gostaria de mudar de emprego, aparentemente.

O leitor notará a semelhança que existe entre certos casos de transposição de advérbio, como em (27a), e a topicalização; com efeito, é bem possível que se trate, nesses casos, do mesmo processo. Mas o advérbio tem outras possibilidades de transposição, exemplificadas em (27b, c, d), que não são cobertas pela topicalização ha-

bitual (esta só é responsável pela transposição para o início do período).

Predeterminante transposto

Os elementos *todos* e *ambos* (que têm a função de **predeterminantes**) podem ocorrer em diversas posições, o que gera grupos de correspondência como o seguinte:

(28) a. Todas as hienas sofrem de gastrite.
b. As hienas todas sofrem de gastrite.
c. As hienas sofrem todas de gastrite.

Um aspecto da relação de correspondência que é importante ressaltar desde já é que nem sempre se trata de uma relação simétrica. Assim, tomando os exemplos (21) e (22), vemos que, para cada frase com elemento topicalizado, existe sempre uma frase com esse elemento não topicalizado; e, para cada frase com elemento (topicalizável) não topicalizado, existe uma outra frase com esse mesmo elemento topicalizado. Por exemplo, qualquer frase composta de **sujeito + verbo + objeto** [como (21)] corresponde a uma outra frase idêntica, mas com a ordenação **objeto + sujeito + verbo** [como (22)]. E vice-versa: qualquer frase composta de **objeto + sujeito + verbo** corresponde a uma outra frase composta de **sujeito + verbo + objeto**. Nesses casos, falaremos de **correspondência total**: (21) e (22) se correspondem totalmente.

Já em outros casos a relação só funciona em uma direção. O exemplo mais conhecido são as duplas de frases ativas e passivas. Sabemos que em muitos casos há duplas como

(29) O rato foi comido pelo gato.
(30) O gato comeu o rato.

Em geral se diz que há "correspondência" entre elas. É verdade que sempre que há uma frase passiva, existe sempre uma frase com os mesmos itens léxicos, mas de forma ativa. No entanto, é muito freqüente haver frases ativas às quais não corresponde nenhuma passiva:

(31) Cristina tinha um pônei castanho.
(32) * Um pônei castanho era tido por Cristina.

Os exemplos são muito numerosos; na verdade, os fatores que governam a possibilidade de passivização são pouco conhecidos. Por ora, diremos simplesmente que a língua não admite passivas em muitos casos. Conseqüentemente, ocorre que, a partir da existência de uma passiva [como (29)], pode-se prever a existência da ativa correspondente, (30). Mas de uma ativa não se pode prever a existência da passiva: há muitas exceções.

Nesses casos, falaremos de **correspondência parcial**; e diremos que as passivas correspondem às ativas, mas as ativas não correspondem às passivas. Ou, ainda, que as passivas são parcialmente correspondentes das ativas (mas as ativas não são correspondentes das passivas de nenhuma forma).

A noção de correspondência não é bem definida na literatura — e, na verdade, não é muito fácil de definir com precisão. No capítulo 7, dou a definição que me parece mais

adequada (sem querer dizer que não apresente problemas). Por outro lado, quase todas as linhas atuais de análise lingüística admitem algum tipo de correspondência; ou seja, aceitam que há um relacionamento sintático especial entre frases como (21) e (22), além do evidente parentesco semântico. Aqui, portanto, também seguiremos essa tendência, admitindo a existência de grupos de correspondência que agregam certas estruturas (frases e também sintagmas menores). A correspondência entre estruturas será, pois, considerada mais um dos "fatos gramaticais" e será utilizada na definição das funções sintáticas propostas no capítulo 3.

No capítulo 7 estudaremos os grupos de correspondência existentes em português, listando-os e definindo as relações formais que vinculam seus membros. Ali também discutiremos a definição formal de "correspondência".

2.2. COMPONENTES DE UMA DESCRIÇÃO GRAMATICAL

2.2.1. Níveis de análise

2.2.1.1. Fonologia, morfologia, sintaxe, semântica

Numa tentativa de equacionar a imensa complexidade da estrutura das línguas, os lingüistas estabeleceram diversos "níveis de análise", definidos pelos vários pontos de vista sob os quais se pode encarar os fenômenos gramaticais. Por exemplo, ao estudar uma frase como

(33) Ana desprezou Ricardo.

pode-se assumir o ponto de vista do estudo da pronúncia. Nesse caso, serão estudadas regras de pronúncia como a que nos obriga a pronunciar o primeiro ***a*** de *Ana* como uma vogal nasal, por ser tônico e estar logo antes de uma consoante nasal (o ***n***); ou a que nos obriga a pronunciar a vogal final de *Ricardo* como um ***u***, e não um ***o*** etc. A esse estudo das regras de pronúncia de uma língua se dá o nome de **fonologia**.

Mas a mesma frase pode ser estudada de outros pontos de vista: por exemplo, descrevendo a constituição interna das palavras. Desse ponto de vista, podemos observar que a palavra *desprezou* é formada de mais de um elemento: a seqüência *desprez-* mais a seqüência *-ou*. A primeira aparece também em outras formas relacionadas, como *desprezo* (tanto o substantivo como a forma verbal), *desprezível*, *desprezamos*, *desprezado* etc.; mais a seqüência *-ou*, que ocorre em outras formas verbais, como *amou*, *desmanchou* etc. Existem também regras que governam a associação dessas partes de palavras (denominadas **morfemas**), e que impedem a formação de palavras como * *despre-*

zi, * *Ricardou* ou * *Ricardível.* O estudo dos morfemas e de suas associações se denomina **morfologia.**

Voltando à frase (33), podemos ainda definir outro ponto de vista, que leva em conta as maneiras como se associam as palavras para formar frases. Assim, podemos observar que existe uma regra pela qual a terminação de *desprezou* de certo modo depende do elemento que se coloca no lugar de *Ana*; tanto é assim que, se em lugar de *Ana*, colocarmos *nós*, *desprezou* terá de se transformar em *desprezamos*. Podemos, além disso, notar que na frase (33) o elemento que governa a forma de *desprezou* ocorre em primeiro lugar na frase, e que modificações no último elemento da frase (*Ricardo*) não afetam a forma de *desprezou*. Uma terceira observação é que existe um pequeno número de palavras que só podem ocorrer no lugar de *Ana*, e outras, em número igual, que só podem ocorrer no lugar de *Ricardo*: *eu*, *nós* etc. só ocorrem no lugar de *Ana*, e *me*, *nos* só no lugar de *Ricardo*. Todas essas observações têm a ver com a estruturação interna da frase, e constituem um estudo denominado **sintaxe.**

Finalmente, podemos levar em conta o significado transmitido por (33). Por exemplo, podemos observar que *Ana* provavelmente designa uma mulher, e *Ricardo* um homem; que a pessoa desprezada é *Ricardo*, e não *Ana*; que o fato de Ana desprezar Ricardo aconteceu no passado, e assim por diante. Traços de significado como esses são, em parte, o resultado da aplicação de certas regras, que integram a **semântica.**

Chamamos a essas disciplinas que se ocupam dos diferentes aspectos das expressões lingüísticas os **componentes** da gramática. Assim, a gramática de uma língua inclui os seguintes componentes: a **fonologia**, a **morfologia**, a **sintaxe** e a **semântica** dessa língua.

É importante notar que esses quatro componentes não esgotam tudo o que se pode estudar a respeito de uma língua. Não tratam, por exemplo, da história das formas lingüísticas, nem do uso das mesmas em diferentes situações sociais, nem do uso feito pelos falantes de seu conhecimento geral do mundo para ajudar a compreender as frases, nem de muitos outros aspectos importantes. A fonologia, a morfologia, a sintaxe e a semântica (mais o léxico: ver a seção seguinte) constituem o estudo da estrutura interna de uma língua — aquilo que a distingue das outras línguas do mundo, e que não decorre diretamente de condições da vida social ou do conhecimento do mundo.

Neste volume abordaremos o estudo da sintaxe e da semântica do português. A sintaxe será tratada nos capítulos 3 a 8, e a semântica nos capítulos 9 a 11.

2.2.1.2. O léxico

Nos capítulos 12 e 13 desta *Gramática*, tratarei das classes de palavras. Estas são colocadas pelas gramáticas tradicionais no capítulo reservado à morfologia, mas não há razões fortes que sustentem essa posição. As classes de palavras pertencem a um componente do estudo da língua que, a rigor, se distingue da gramática propriamente dita, a saber, o **léxico**. Estudaremos o léxico no capítulo 14.

Por ora, baste-nos saber que no léxico se colocam as informações que não se podem reduzir a regras gerais — ou seja, as informações **idiossincráticas**. Por exemplo: a pronúncia de cada palavra é em parte previsível através de regras da língua (nunca podemos terminar uma palavra com **p**, nem começar com o **r** "brando" de *cara*); mas há também uma parte que precisa ser aprendida caso por caso. Assim, se sabemos que uma palavra começa com **me-**, nem por isso sabemos como continua: pode ser *melado, melaço, melancolia* ou *meu*. Ou, para tomar um exemplo morfossintático, sabermos que *mão* é um substantivo feminino não nos ajuda a saber se *não, cão* e *dão* são também substantivos femininos. Toda essa informação é armazenada no léxico, uma lista de dezenas de milhares de itens. Aprender uma língua não é apenas aprender suas regras, mas ainda memorizar seu léxico (ou grande parte dele).

2.2.2. Como se articulam os componentes

Os componentes da gramática se articulam para, conjuntamente, definir quais são as seqüências que constituem frases corretas da língua e quais as que não constituem. Assim, a frase

(33) Ana desprezou Ricardo.

está construída em conformidade com todas as regras do português, e por isso é uma frase gramaticalmente bem formada; o resultado é que os falantes a aceitam como uma frase legítima do português.

Agora vejamos a seqüência

(34) * Ana desprezeu Ricardo.

Essa seqüência não é uma frase bem formada porque, ao construí-la, desobedecemos a uma regra morfológica, a que estipula que *desprezar* não pode ter a forma *desprezeu* (ou seja, é um verbo da primeira conjugação, e não da segunda). Correspondentemente, os falantes a rejeitam.

Na seqüência

(35) * Ricardo comprou um livro amarelas.

foi desobedecida uma regra sintática que exige concordância de gênero e número entre um substantivo (aqui, *livro*) e o adjetivo que o acompanha (*amarelas*). Novamente, o resultado é que a frase é rejeitada.

Finalmente, podemos ter uma frase inaceitável por motivos semânticos, como

(36) * Esse pastel desprezou Ricardo.

A frase é mal formada, mas isso não vem propriamente de sua estruturação morfológica ou sintática; ela simplesmente transmite um conteúdo implausível, estranho, já que pastéis não "desprezam" ninguém. A frase (36) só poderia ser utilizada em algum contexto particular, como em uma fábula, por exemplo.

Conforme se vê, é como se uma frase, para ser bem formada (e portanto aceitável), tivesse que passar por um crivo múltiplo, uma espécie de seção de controle de qualidade. Só serão bem formadas as frases que passarem com sucesso por todos os testes — isto é, as que forem construídas sem desobediência a regras fonológicas, morfológicas, sintáticas e semânticas. Em outras palavras, são bem formadas as frases construídas de acordo com a gramática da língua.

Aqui é preciso fazer uma observação importante: não se entenda do que se disse acima que um livro como este autoriza os falantes a usarem esta ou aquela construção. Muito menos se deve entender que os gramáticos têm a prerrogativa de autorizar ou desautorizar formas e construções da língua. A língua tem uma vida e um funcionamento próprios, que não dependem em absoluto dos desejos, crenças ou determinações dos gramáticos ou lingüistas. Quando dizemos que uma frase precisa estar de acordo com a gramática da língua para ser bem formada, queremos dizer que cada falante nativo tem em sua cabeça, como que programado, um conjunto de instruções para construir as formas de sua língua. Essas instruções foram adquiridas na infância e, na verdade, de certa forma foram deduzidas pelo próprio falante; trata-se de uma parte do conhecimento do mundo que cada pessoa carrega em si. O papel da escola, do professor e dos livros denominados "gramáticas" nesse aprendizado é muito reduzido; aprende-se a língua materna de maneira informal (mas extremamente eficiente).

Aqui não é o lugar para discutir a questão da gramática como conjunto de instruções programadas no cérebro. Fique claro, no entanto, que isso é o que a gramática é em primeiro lugar; o que o livro chamado "gramática" pode fazer é tentar representar esse conhecimento através de abstrações, tais como regras, classes, princípios, definições etc. O que autoriza ou desautoriza as formas da língua (palavras, frases, sintagmas) é a gramática que todos levamos impressa em nosso cérebro, não a gramática enquanto livro ou instituição educacional.

Assim, deve-se entender a gramática exposta neste livro como um conjunto de instruções que o falante da língua domina implicitamente — ele sabe muito bem pô-las em ação,

ao julgar a boa ou má formação de uma frase ou de uma palavra. Mas isso não quer dizer que ele tenha consciência dessas instruções, não mais do que tem consciência dos processos de sua digestão ou circulação. É um mecanismo que ele põe em funcionamento de maneira automática; a tarefa do lingüista é tentar explicitar esse mecanismo e estudá-lo. O resultado dessa tentativa de explicitação é uma **gramática** tal como apresentada nos livros usualmente chamados "gramáticas", inclusive este.

Os componentes estudados neste volume, a sintaxe e a semântica, diferem em um ponto importante. A sintaxe tem como única função definir quais são as frases bem formadas na língua; assim, trata-se apenas de um conjunto de instruções sobre o modo de construir frases em português. A semântica também tem uma função paralela, ou seja, um conjunto de instruções sobre o modo de construir frases que tenham sentido; mas a semântica, além disso, é um dos pontos de contato da língua com o mundo exterior. As regras semânticas, além de atuarem como filtros, excluindo as frases semanticamente mal formadas, também atribuem a cada construção um significado, denominado **interpretação semântica.**

Para dar um exemplo, seja a frase

(37) Um velho professor atropelou meu cachorro.

As regras sintáticas estabelecem que *atropelou* deve concordar com *um velho professor*, que *meu* deve concordar com *cachorro*, que *um velho professor* (sujeito) deve vir antes de *atropelou* (núcleo do predicado) etc. Uma vez esgotadas todas as condições semelhantes, está concluída a análise sintática da frase.

As regras semânticas atribuem interpretações semânticas a (37) e a suas partes. Assim, elas devem explicitar que *velho* se refere a *professor*, e não a *cachorro* (isto é, o professor é velho, e não o cachorro); devem explicitar que *um velho professor* praticou a ação denotada por *atropelou* e que *meu cachorro* sofreu essa mesma ação; devem ainda deixar claro que o professor é provavelmente desconhecido do ouvinte; e assim por diante. Uma vez feitas todas as atribuições de significado, temos a interpretação semântica completa da sentença.

Mas essa interpretação ainda tem a função de filtro, ou seja, será necessário excluir as frases (ou partes de frases) que não façam sentido. Existem condições especiais que estipulam o que é que "faz sentido" (essas condições podem ser consideradas extralingüísticas). Por exemplo, se na frase *escreveu* estivesse no lugar de *atropelou*, a frase seria marcada como semanticamente mal formada, porque "escrever um cachorro" não faz sentido.

Como se vê, a semântica tem duas funções na definição das expressões bem formadas, e a sintaxe uma só. Veremos muitos exemplos

da atuação das regras sintáticas nos capítulos 3 a 8 e das regras semânticas nos capítulos 9 a 11. Por ora, vou apenas concluir esta seção com um apanhado final do que vem a ser uma gramática como esta.

Um livro de gramática tem como objetivo fundamental descrever o sistema de conhecimentos que os falantes de uma língua possuem e que lhes permite reconhecer as expressões bem formadas de sua língua. Uma gramática bem sucedida é formada por um conjunto de instruções sobre como construir as formas da língua (palavras, sintagmas e orações), mais um conjunto de regras semânticas que atribuem a essas formas determinados significados.

2.3. DEFINIÇÕES E REGRAS

Vejamos finalmente um aspecto importante, mas muitas vezes negligenciado, de qualquer estudo gramatical: a necessidade de propor regras e definições explícitas.

A situação das definições na gramática tradicional é tal que algumas pessoas chegam a negar a possibilidade de se definir uma série de noções fundamentais. Assim, dirão: "O substantivo não se define; aprende-se a reconhecer". Isso é certamente verdade no que diz respeito à aprendizagem prática da língua: o falante aprende a utilizar corretamente os substantivos (por exemplo, nunca tenta conjugá-los como se fossem verbos), sem nunca ter consciência de alguma definição explícita.

Mas o gramático, ou o estudioso de gramática, não é um falante comum; seu objetivo não é aprender a usar a língua, mas pesquisar seu funcionamento interno. Para ele, é importante estabelecer definições, pois estas explicitam parte do conhecimento implícito que o falante adquire e que chamamos a estrutura da língua. A gramática, enquanto explicitação da estrutura da língua, não decorre da aprendizagem da mesma; precisa ser investigada e formulada por especialistas. Todas as pessoas respiram e não precisam para isso saber que têm pulmões, traquéia, brônquios etc. Mas o estudo científico da respiração envolve a explicitação da anatomia e da fisiologia do aparelho respiratório. Ninguém sustentaria que essa explicitação é desnecessária, porque "as pessoas aprendem a respirar sem saber nada disso"; reconhece-se simplesmente a diferença fundamental entre estudar a respiração e respirar. A atitude frente à estrutura da língua deve ser exatamente a mesma: para o usuário, o conhecimento explícito não é relevante, mas para o estudioso é.

A estrutura da língua se explicita através de instrumentos tais como definições, regras etc. A seguir vou fazer algumas considerações (que para muitos serão óbvias) sobre as qualidades que esses instrumentos devem ter. Vou partir do exemplo

das definições, mas a argumentação vale igualmente para quaisquer afirmações gramaticais.

Para que serve uma definição? A formulação de uma definição só se justifica se ela possibilita a identificação de uma entidade gramatical. É inútil uma definição que só possa ser aplicada corretamente por quem já saiba o resultado de antemão. Assim, uma boa definição de "sujeito" deve possibilitar a identificação do sujeito de uma oração por alguém que não saiba qual é o sujeito, ou mesmo o que é um sujeito — simplesmente aplicando a definição. Para que uma definição atinja tais objetivos, é necessário que ela tenha (pelo menos) as qualidades seguintes:

(a) ser explícita: isto é, fornecer todos os elementos necessários à sua aplicação;
(b) ser adequada ao usuário em perspectiva: isto é, utilizar noções e termos que o usuário já conheça;
(c) ser adequada à realidade lingüística: isto é, descrever fatos reais da língua, e não fantasias ou opiniões pessoais a respeito dela.

Vamos examinar algumas definições e outras afirmações gramaticais, levando em conta essas exigências. Primeiramente, podemos considerar esta definição de "substantivo":

> Substantivos são palavras que designam os seres. [Cegalla, 1987, p. 110]

É claro que só podemos aplicar essa definição se já sabemos bem claramente o que é uma palavra; o que é um ser; e o que quer dizer *designar*. Quanto a palavras e designações, creio que seria possível dar uma conceituação toleravelmente precisa. Mas o que dizer da noção de "ser"? O que, exatamente, é um "ser"? Todos concordam que *gato* designa um ser, assim como *menino*. Mas que dizer de *lingüística, pôr-de-sol, universo, teoria* e *sabonete*? Aqui o julgamento das pessoas vai certamente variar; e acredito que em certos casos, como o de *teoria*, a maioria negará que se trata de um "ser". No entanto, *teoria* é um substantivo de acordo com a classificação tradicional.

O que temos aqui é uma definição que deixa de atender à exigência (b) acima: lança mão de uma noção (a de "ser") que não é dominada pelos usuários (nem, talvez, por ninguém).

Para fazer com que a definição acima funcione, poderemos procurar uma saída mais ou menos assim: diremos que *teoria* é um substantivo porque pode ser núcleo do sujeito de uma oração. Mas nesse caso estaremos violando a exigência (a): estamos utilizando, para identificar o substantivo, um elemento que não figura na definição (pois esta não menciona a possibilidade de ele ser núcleo de um sujeito). Se a definição de sujeito depende desse elemento para funcionar, não será explícita, o que a tornará inaplicável na prática.

A definição de "substantivo" dada acima, portanto, é inadequada aos usuários; e, além do mais, é em geral não-explícita, em vista da classificação das palavras que encontramos feita na prática.

Vejamos um caso de definição inadequada à realidade lingüística. Define-se às vezes o uso do modo subjuntivo da seguinte forma:

> Emprega-se o modo subjuntivo para exprimir um fato possível, incerto, hipotético, irreal ou dependente de outro. [Cegalla, 1987, p. 487]

No entanto, encontramos o subjuntivo em frases como

(38) É muito triste que Cacá esteja desempregado.

Acontece que quem ouve (38) recebe a mensagem de que Cacá **está** de fato desempregado — não há nada de hipotético, nem incerto, nem (apenas) possível, nem irreal a respeito disso: Cacá está desempregado, e isso é muito triste. Da mesma forma, não se pode dizer que o fato de Cacá estar desempregado seja "dependente" de isso ser muito triste.

Como se vê, a afirmação sobre o subjuntivo (que deveria ter como uma de suas funções possibilitar ao estudioso a identificação do subjuntivo) não respeita os fatos da língua. Uma frase como (38) é normal em português, mas não há meio de encaixá-la na definição dada.

Em conclusão, as definições (assim como as demais afirmações gramaticais) são necessárias para que se possa elaborar a gramática da língua; mas é fundamental que se submetam a certas exigências (como as que vimos acima), pois de outro modo não estaremos explicitando a estrutura da língua. Esta *Gramática* foi elaborada com a intenção de escapar a esses defeitos; um grande esforço foi envidado no sentido de fornecer definições, regras etc. que sejam realmente aplicáveis, que descrevam as realidades da língua, e que constituam em seu conjunto um todo coerente e logicamente estruturado. Meu sucesso nessa tentativa, certamente, não deve ter sido total (longe disso). Mas é nessa direção que nos devemos mover; o leitor é convidado a criticar cada uma das afirmações aqui encontradas, e rejeitá-la se for o caso. Só assim haverá progresso verdadeiro no estudo da gramática portuguesa.

2.4. SITUAÇÕES ANAFÓRICAS E NÃO-ANAFÓRICAS

Uma última observação importante a ser feita a respeito do estudo da gramática tem a ver com a distinção entre **situações anafóricas** e **situações não-anafóricas**.

A gramática estuda a estrutura das formas lingüísticas e seu significado dentro dos limites máximos do período, isto é, sem considerar a even-

tual (e habitual) inserção dos períodos dentro de contextos lingüísticos e extralingüísticos. Isso não quer dizer que se negue a relevância do estudo do contexto, é claro; apenas, defende-se a posição de que o estudo do período é parte do estudo dos enunciados. O contexto é estudado por disciplinas especiais: a **análise do discurso** e a **pragmática**; a primeira se ocupa da estrutura e interpretação de textos, e a segunda da relação dos enunciados lingüísticos com a situação extralingüística em que se inserem.

Vimos que o período é o limite máximo dos estudos estritamente gramaticais, e isso lhes impõe certas limitações. A principal destas é justamente a limitação de estudar a estrutura e a interpretação das formas em **situações não-anafóricas**. Essa limitação não é absoluta, mas tem de ser levada em conta em certos pontos importantes.

Chamamos "situação anafórica" aquela em que uma estrutura está reduzida porque ocorre na vizinhança de outra estrutura de certo modo paralela; os elementos comuns às duas estruturas são total ou parcialmente omitidos em uma delas. Um exemplo é o diálogo

(39) A: Você fez o exercício?
B: Fiz.

O constituinte *o exercício* é omitido na resposta de B, que se interpreta como significando "fiz o exercício". Dizemos então que *fiz* na resposta está em situação anafórica.

Agora: o estudo gramatical se faz geralmente considerando apenas situações não-anafóricas. Assim, como veremos no capítulo 6, afirma-se que o verbo *fazer* exige a presença de um objeto direto (como *o exercício*). Estritamente falando, isso não é verdadeiro, como se viu acima; mas a omissão do objeto direto só se dá em situações anafóricas e, portanto, não conta para efeitos de descrição gramatical.

Outro exemplo é a afirmação de que a palavra *este* não pode ocorrer sozinha como sujeito de uma oração. Essa afirmação só vale para situações não-anafóricas, porque podemos ter

(40) Carolina quer aquela bicicleta, mas esta é bem melhor.

A afirmação mencionada não é invalidada por esse exemplo, porque *esta* se encontra em situação anafórica e não conta para efeitos de estudo sintático.

As situações anafóricas são objeto de estudo gramatical em certos casos particulares, cuidadosamente delimitados; alguns exemplos serão vistos no capítulo 11. Na maioria dos casos, são consideradas como fazendo parte do campo de interesse da análise do discurso.

SEGUNDA PARTE: SINTAXE

3 A oração simples

3.1. A FRASE E A ORAÇÃO

3.1.1. Frase, oração, período

O termo **frase** é utilizado de maneira geral para designar uma unidade do discurso bastante difícil de definir. A conceituação oferecida por Mattoso Camara é provavelmente a melhor, embora não chegue a ser uma definição plenamente satisfatória:

> Unidade de comunicação lingüística, caracterizada [...] do ponto de vista comunicativo — por ter um propósito definido e ser suficiente para defini-lo —, e do ponto de vista fonético — por uma entoação [...] que lhe assinala nitidamente o começo e o fim. [Camara, 1977, p. 122]

Poderíamos acrescentar que, na escrita, a frase é delimitada por uma maiúscula no início e por certos sinais de pontuação (. ! ? ...) no final.

Essa definição apresenta problemas, que não serão discutidos aqui. Baste-nos reconhecer que geralmente é possível identificar frases, embora as bases para essa identificação permaneçam em parte obscuras. Assim, são frases os enunciados seguintes:

(1) Meu livro tem mais de mil páginas.
(2) Quantas páginas tem o seu livro?
(3) Vá à padaria e traga oito pãezinhos.
(4) Você poderia me trazer um pãozinho?
(5) Que calor!
(6) Quantas páginas?
(7) Mas que livro enorme!

Oração é uma frase que apresenta determinado tipo de estrutura interna, incluindo sempre um **predicado** e freqüentemente um **sujeito**, assim como vários outros termos. "Predicado" e "sujeito" serão conceituados na seção 3.2.; por ora, observarei apenas que as frases de (1) a (4) são orações (às vezes compostas, por sua vez, de mais de uma oração); as frases (5) a (7) não são orações, por carecerem de predicado.

As frases não-oracionais nem por isso deixam de ter estrutura

analisável; em geral, verifica-se que se compõem de elementos que também ocorrem dentro de orações — ou seja, são como que fragmentos de orações. As frases não-oracionais estão muito pouco estudadas, de maneira que não será possível darlhes a devida atenção nesta *Gramática*. Isso não significa, claro, que não sejam interessantes; em particular, a hipótese de que uma frase não-oracional é sempre composta de um fragmento de oração merece ser investigada.

Tradicionalmente, emprega-se também a designação **período** para as orações que constituem uma frase. Assim, em

(3) Vá à padaria e traga oito pãezinhos.

há duas orações, a saber: (a) *vá à padaria*; (b) *traga oito pãezinhos*. Além disso, há ainda uma terceira oração, que compreende as duas mencionadas, mais a palavra *e*, ou seja, essa terceira oração é a íntegra de (3). Como se vê, a terceira oração é coextensiva com a própria frase e seria, portanto, um período. Não vejo inconveniente nessa nomenclatura, desde que fique claro que **um período é sempre uma oração** (discuto mais cuidadosamente esse ponto nas seções 5.2.2.2. e 5.2.2.3.). Naturalmente, nem toda oração é um período, já que muitas orações não são coextensivas com a frase de que fazem parte; por exemplo, *vá à padaria* em (3) é uma oração, mas não um período.

A **sintaxe** é a parte da gramática que estuda as orações e suas partes — ou seja, a estrutura interna da oração.

3.1.2. Força ilocucionária

Voltemos a alguns dos exemplos de frases dados na seção precedente:

(1) Meu livro tem mais de mil páginas.
(2) Quantas páginas tem o seu livro?

Essas duas frases, além de mostrarem certas diferenças de estrutura, têm funções claramente diferentes no discurso. A primeira é usada, normalmente, para fazer uma declaração: o falante informa algo a respeito de seu livro. Já a segunda não traz informação; antes, solicita uma informação: o falante admite, de saída, que seu interlocutor tem um livro, e quer saber quantas páginas tem esse livro. Dizemos, então, que a frase (2) tem uma **força ilocucionária** diferente da da frase (1): (2) é uma **pergunta**, ao passo que (1) é uma **declaração**.

Os tipos de força ilocucionária que encontramos geralmente no discurso são, entre outros, os seguintes:

— **declaração** [exemplo: (1)];
— **pergunta** [exemplo: (2)];
— **exclamação** [exemplo: (3)];
— **ordem** [exemplo: (4)];
— **pedido** [exemplo: (5)];

— **promessa**;
— **expressão de um desejo** etc.

É preciso observar, antes de mais nada, que a força ilocucionária não é uma propriedade das frases propriamente ditas, mas das frases **em determinados contextos**. Assim, a frase

(3) Vá à padaria e traga oito pãezinhos.

pode ser, segundo a situação, uma ordem ou um pedido. A frase

(4) Você poderia me trazer um pãozinho?

seria normalmente interpretada como um pedido. Mas sua forma nos autoriza a entendê-la como uma pergunta: digamos que duas pessoas estão imaginando uma situação hipotética. Uma delas diz:

(8) Suponha que você fosse um empregado desta padaria, e eu um mendigo que lhe pedisse uma esmola. Você poderia me trazer um pãozinho?

A resposta normal não seria trazer realmente o pãozinho, mas responder algo como:

(9) Não, porque o dono nunca deixaria.

Como se vê, a estrutura da frase não é suficiente para determinar sua força ilocucionária; muitas vezes é preciso ter também informações sobre o contexto em que a frase é usada.

Isso não quer dizer que a estrutura seja irrelevante. Por exemplo, é muito difícil imaginar um contexto em que a frase

(3) Vá à padaria e traga oito pãezinhos.

seja entendida como uma pergunta, ou como uma declaração, ou como uma exclamação. Derivamos nossa percepção da força ilocucionária a partir da estrutura da frase, **mais** outros fatores do contexto lingüístico e situacional.

O que nos interessa aqui é que há certas estruturas sentenciais que se especializam em veicular determinado tipo de força ilocucionária. Assim, falamos de **frases interrogativas**, que em geral se usam para exprimir perguntas. Como vimos, isso não quer dizer que elas **só possam** exprimir perguntas; mas "pergunta" faz parte, digamos, de seu significado básico, literal (ver a seção 9.1. para mais discussão sobre a questão do significado literal). O que é importante para nós no momento é não confundir uma **frase** (ou **oração**) **interrogativa** com uma **pergunta**; uma frase interrogativa é uma frase que apresenta determinado tipo de estrutura, ao passo que uma pergunta é uma frase (qualquer que seja sua estrutura) que é usada, em determinado contexto, com a força ilocucionária de um pedido de informação.

Na seção seguinte darei os tipos de estrutura que as orações portuguesas apresentam, relacionando-as com o tipo de força ilocucionária

que cada uma tipicamente (mas não obrigatoriamente) veicula.

3.1.3. Tipos de orações

As orações do português podem ser distinguidas em cinco tipos: **imperativas**, **interrogativas**, **exclamativas**, **declarativas** e **optativas**, estas últimas de importância secundária. Destas, as declarativas apresentam uma grande variedade de estruturas e são consideradas o caso comum, não marcado. Assim, vamos primeiro ver como se caracterizam do ponto de vista formal os primeiros três tipos; as declarativas serão então definidas como as orações que não se encaixam em nenhum dos outros tipos.

3.1.3.1. Imperativas

As **orações imperativas** se caracterizam por apresentarem uma forma especializada do verbo, denominada **imperativo**. Assim, por exemplo, em

(10) Frite um ovo para o doutor.

a forma *frite* é um imperativo (do verbo *fritar*) e, por isso, a oração é chamada de imperativa.

Outra característica das orações imperativas é que geralmente aparecem sem sujeito [como (10)]. Mas essa característica não é suficiente para defini-las, porque, como veremos, as orações não-imperativas podem também ocorrer sem sujeito; e, além disso, às vezes encontramos imperativas com sujeito:

(11) Você frite um ovo para o doutor; e vocês peguem o uísque no armário.

Isso se dá, em geral, quando duas orações imperativas são colocadas em contraste, como no exemplo (11).

As orações imperativas são tipicamente utilizadas para veicular **ordens** e **pedidos**.

3.1.3.2. Interrogativas

As **orações interrogativas** compreendem dois subtipos principais: **interrogativas abertas** (também chamadas **interrogativas de "Q"**) e **interrogativas fechadas** (ou **interrogativas de sim-ou-não**).

As **interrogativas abertas** se caracterizam por conterem um elemento interrogativo, tirado da lista seguinte:

(o) que, (o) quê, quem, quando, como, por que, onde, qual.

O elemento interrogativo pode ser colocado no início da oração, mas isso não é obrigatório. Além disso, as orações interrogativas são (quando não subordinadas) seguidas de ponto de interrogação. Exemplos:

(12) Quem a Renata vai escolher para padrinho?
(13) Você vai fazer o quê?
(14) Você prefere que raça de cachorro?
(15) Quando é que a sua mãe faz pão de queijo?

O elemento interrogativo tem uma função dentro da oração: em (12), por exemplo, *quem* é objeto direto.

As **interrogativas fechadas** são geralmente marcadas apenas pelo ponto de interrogação:

(16) Sua mãe vai fazer pão de queijo hoje?

É bom observar que, embora tanto as interrogativas abertas quanto as fechadas ocorram na escrita com ponto de interrogação, na fala suas marcas entonacionais são diferentes. As interrogativas abertas têm entoação descendente, idêntica à de certas declarativas; já as interrogativas fechadas têm entoação final ascendente.

Além desses dois tipos principais de interrogativas, pode-se mencionar um terceiro tipo, algo marginal, chamado **interrogativa-eco**. Na escrita, as interrogativas-eco geralmente não se distinguem das abertas, mas na fala o elemento interrogativo recebe um tipo especial de entoação, alta e ascendente, como em

(17) Sua mãe vai fazer O QUÊ?

Nas interrogativas-eco, o elemento interrogativo nunca é colocado no início da oração, a não ser que se trate do sujeito; essa característica distingue em parte as interrogativas-eco das abertas, mesmo na escrita.

Em alguns casos, não muito bem conhecidos, pode ocorrer inversão de posição do sujeito nas interrogativas abertas (o sujeito é colocado depois do verbo); isso ocorre quando o interrogativo não é sujeito:

(18) Quando chegou o avião?
(19) Como vai o seu pai?
(20) Onde está o cachorro?
(21) Quem é você?
(22) Que é isso?

Em outros casos, a inversão dá resultado marginal, ou mesmo inaceitável:

(23) ?? Por que trabalha ele tanto?
(24) ?? Como conseguiu ele esse emprego?

(Estou usando "??" como marca de aceitabilidade baixa, mas não tão baixa quanto a sinalizada por "*".)

A possibilidade de inversão parece ter algo a ver com o verbo: é sempre mais aceitável com *ser* e *estar* do que com os outros verbos; e talvez também tenha a ver com o elemento interrogativo: com *por que* a inversão nunca é plenamente aceitável. Este é um tema que ainda falta investigar.

Com interrogativas fechadas, a inversão se sente hoje como arcaica:

(25) ?? Está seu irmão em casa?
(26) ?? Trabalham vocês aqui?

As orações interrogativas servem, tipicamente, para veicular perguntas (solicitação de informação). No caso das interrogativas-eco, essa solicitação sugere que o falante já recebeu a informação, mas não a entendeu bem, ou então está algo incrédulo a respeito dela.

3.1.3.3. Exclamativas

As **orações exclamativas** têm estrutura semelhante à das interrogativas, mas não são marcadas com ponto de interrogação; na fala, quando semelhantes às interrogativas fechadas, as exclamativas não apresentam a entoação final ascendente que caracteriza aquelas interrogativas:

(27) Como você é incompetente!
(28) Que roupa você foi vestir!
(29) Elas foram embora!

Do ponto de vista sintático, portanto, as exclamativas se distinguem muito pouco das interrogativas e não apresentam peculiaridades estruturais de interesse. São seguidas, na escrita, de ponto de exclamação (!) e, na fala, mostram um tipo particular de contorno entonacional.

Do ponto de vista da força ilocucionária, pode-se dizer que as exclamativas expressam exclamação e surpresa, muitas vezes [como em (29)] simultaneamente com uma declaração.

3.1.3.4. Declarativas

As **orações declarativas**, como vimos, se definem negativamente, por não apresentarem os traços distintivos das imperativas (verbo no imperativo), das interrogativas (elemento interrogativo e ponto de interrogação) e das exclamativas (ponto de exclamação). As declarativas podem apresentar estruturas bem variadas, como mostram os exemplos abaixo:

(30) O seu bode comeu minha camisa.
(31) Minha camisa foi comida pelo seu bode.
(32) Minha camisa, o seu bode comeu.
(33) Foi o seu bode que comeu minha camisa.

Tipicamente, as declarativas são usadas para expressar declarações. Mas, como vimos, também podem expressar outros tipos de força ilocucionária, dependendo do contexto lingüístico e/ou extralingüístico. Por exemplo:

(34) Você poderia me ajudar aqui. (**pedido**)
(35) Eu ainda não sei o seu nome. (**pergunta**)
(36) Eu gostaria que alguém me ajudasse. (**desejo ou pedido**)

3.1.3.5. Optativas

Um tipo de oração de importância bem menor, por só ocorrer raramente e em casos especializados, são as **orações optativas**. Caracterizam-se, morfossintaticamente, por terem verbo no subjuntivo na oração principal (fora das optativas, o subjuntivo só ocorre em subordinadas). Também são freqüentemente introduzidas pelo exclamativo *que*, podendo ser marcadas com ponto de exclamação. Quanto ao significado, expressam desejo:

(37) Deus me ajude!
(38) Que a sorte o acompanhe!

As optativas não iniciadas por *que* também podem apresentar inversão sujeito/verbo:

(39) Queira Deus que você chegue vivo a São Paulo!
(40) Possa essa idéia ser aceita por todos.

Aqui, novamente, não são bem conhecidos os fatores que possibilitam a inversão.

3.1.3.6. Sumário

As orações se dividem, como acabamos de ver, em diversos tipos, segundo sua estrutura formal: imperativas, interrogativas (abertas, fechadas e eco), exclamativas, declarativas e optativas.

É preciso enfatizar que essa classificação é formal, morfossintática, e só diz respeito ao tipo de estrutura que cada uma manifesta. Existe, é verdade, alguma associação de cada tipo com um ou mais tipos de força ilocucionária; por exemplo, as interrogativas tipicamente expressam perguntas, as imperativas pedidos ou ordens etc. Mas os dois planos devem ficar estritamente separados: **oração interrogativa** não é o mesmo que **pergunta**. A primeira dessas categorias pertence à sintaxe, a segunda à semântica (ou, talvez melhor, à pragmática).

Nos capítulos 3, 4 e 5 estudaremos detalhadamente o que se sabe da estrutura formal das orações. Esse estudo, como vimos, se denomina **sintaxe**. O estudo dos tipos de força ilocucionária está além dos objetivos deste livro; deveremos contentar-nos, portanto, com a rápida noção dada acima.

3.2. FUNÇÕES SINTÁTICAS

3.2.1. Estrutura interna da oração

3.2.1.1. Definição formal das funções

Vamos começar a estudar agora as orações em detalhe. Como se verá, elas apresentam uma estrutura interna muito rica e complexa; o es-

tudo dessa estrutura recebe o nome tradicional de *análise sintática.*

A análise sintática que passarei a expor difere em muitos aspectos importantes da análise sintática encontrada nas gramáticas usuais. Em primeiro lugar e acima de tudo, é uma análise *sintática* e não *semântica* — ou seja, uma descrição da estrutura *formal* da oração, sem levar em conta o eventual significado dela e de seus componentes. Assim, definirei adiante o "sujeito" como sendo "o termo com o qual o verbo concorda", exprimindo desse modo um aspecto da organização formal da oração; isso sem dizer nada acerca do que o sujeito tipicamente significa dentro da oração — seja "o elemento que pratica a ação", seja "o elemento sobre o qual se faz uma declaração", seja, enfim, qualquer outra noção semântica.

Isso não quer dizer, evidentemente, que o estudo do significado dos diversos termos seja impossível ou sem interesse. Mas os pressupostos da presente análise exigem que sejam separados, para efeitos descritivos, os fatos formais e os fatos semânticos (ver a seção 2.1., acima; repito que neste particular não se faz mais aqui que seguir a orientação predominante da lingüística moderna). Na terceira parte desta *Gramática,* teremos ocasião de examinar as relações que se podem observar entre as funções sintáticas e a interpretação semântica das orações. Ali será possível considerar perguntas como: "Qual é o significado típico do sujeito da oração?" Como veremos, a resposta é geralmente muito mais complexa do que dão a entender as gramáticas tradicionais.

3.2.1.2. A hierarquia dos constituintes

A oração se estrutura de maneira hierárquica, isto é, contém constituintes que, por sua vez, contêm outros constituintes. É preciso levar esse fato em conta ao se fazer a análise. Por exemplo, digamos que se vai analisar a oração

(41) Meus vizinhos arranjaram um cachorro horrivelmente barulhento.

Podemos fazer um primeiro corte, definindo os grandes constituintes (ou **sintagmas**) da oração, da seguinte forma:

(42) [Meus vizinhos] — [arranjaram] — [um cachorro horrivelmente barulhento]

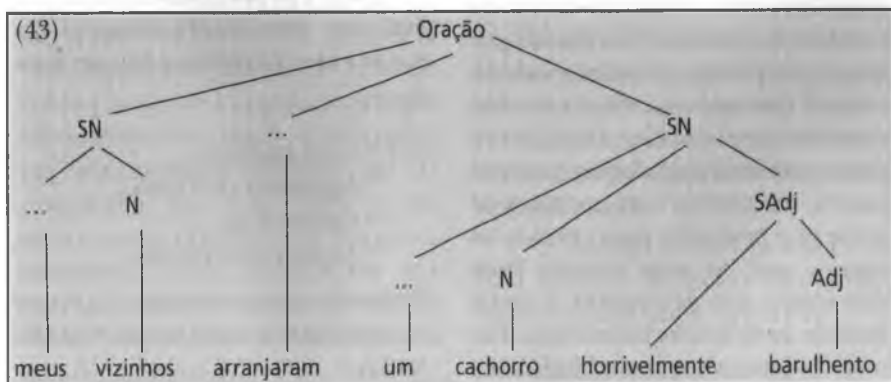
Esses são os **constituintes imediatos** da oração; cada um deles terá uma função especial (essas funções se denominam, na ordem: "sujeito", "predicado" e "objeto direto").

Mas é fácil verificar que alguns desses constituintes têm, por sua vez, uma estrutura interna sintaticamente caracterizável. Assim, *meus vizinhos* se divide em *meus* e *vizinhos,* e cada

uma dessas palavras tem sua função sintática dentro do sintagma *meus vizinhos* (chamo a essas funções, respectivamente, "possessivo" e "núcleo do sintagma nominal"). O constituinte *arranjaram* é sintaticamente simples, por ser formado de uma única palavra, e portanto não pode ser mais analisado em termos sintáticos (pode sê-lo em termos morfológicos). O constituinte *um cachorro horrivelmente barulhento*, por sua vez, é complexo, e precisa ser analisado. Conforme veremos, ele se analisa em um "determinante" (*um*), um "núcleo do sintagma nominal" (*cachorro*) e um "modificador" (*horrivelmente barulhento*). Esses três termos são os constituintes imediatos do sintagma nominal *um cachorro horrivelmente barulhento*. Finalmente, o constituinte que funciona como modificador desse sintagma, *horrivelmente barulhento*, ainda pode ser analisado sintaticamente; proporei adiante que se divida em um "intensificador" (*horrivelmente*) e um "núcleo do sintagma adjetivo" (*barulhento*).

Observe-se agora a hierarquia na estruturação da oração: o termo *meus vizinhos* é o sujeito da oração; mas o termo *meus* não é o possessivo **da oração**; é o possessivo do sintagma nominal *meus vizinhos*. *Meus* só é parte da oração indiretamente, porque faz parte de um sintagma que por sua vez faz parte da oração. Da mesma forma, *horrivelmente* não é intensificador da oração, nem do sintagma nominal *um cachorro horrivelmente barulhento*; é o intensificador do sintagma adjetivo *horrivelmente barulhento*. Esse sintagma adjetivo faz parte do sintagma nominal *um cachorro horrivelmente barulhento*, e esse sintagma nominal faz parte da oração (41).

Essa estruturação hierárquica é muitas vezes apresentada em forma de árvore, conforme abaixo.

Alguns nódulos da árvore estão sem nome; isso não tem importância no momento, pois só nos interessa apresentar visualmente a estruturação hierárquica dos constituintes da oração.

3.2.1.3. Diferenças de comportamento gramatical

O que é que faz com que todos esses termos tenham "funções" diferentes? Por que não poderíamos, por exemplo, dizer que *meus vizinhos* tem a mesma função na oração que *um cachorro horrivelmente barulhento*? Por que *meus* é "possessivo", ao passo que *um* é "determinante"?

Essas perguntas serão devidamente respondidas no decorrer deste capítulo, à medida que formos conceituando e definindo as diferentes funções. Mas é possível adiantar alguma coisa desde já: cada constituinte tem uma função de acordo com seu comportamento gramatical. E o comportamento gramatical se define em termos dos tipos de "fatos gramaticais" vistos na seção 2.1.: sua posição na seqüência de constituintes, suas relações de regência ou concordância com outros elementos, suas possibilidades de retomada pronominal etc.

Assim, temos de atribuir funções diferentes a *meus vizinhos* e a *um cachorro horrivelmente barulhento* por uma razão principal: somente *meus vizinhos* é que está em relação de concordância com o verbo. Assim, *arranjaram* está em uma forma especial para se harmonizar com os traços de número e pessoa de *meus vizinhos*; sabemos que, se *meus vizinhos* fosse substituído por *meu vizinho*, o verbo teria de assumir a forma *arranjou*. Por outro lado, nenhuma modificação de número ou pessoa no sintagma *um cachorro horrivelmente barulhento* poderia acarretar a necessidade de mudanças na forma do verbo. Em outras palavras, o comportamento gramatical do sintagma *meus vizinhos* na frase (41) é nitidamente diferente do comportamento gramatical do sintagma *um cachorro horrivelmente barulhento*. Por isso, atribuímos a cada um deles uma função diferente; aqui, seguindo a tradição, denomino "sujeito" a função de *meus vizinhos*, e "objeto direto" a de *um cachorro horrivelmente barulhento*.

O mesmo se dá com a diferença de função entre *meus* de *meus vizinhos* e *um* de *um cachorro horrivelmente barulhento*. A diferença entre o possessivo e o determinante será devidamente examinada na seção 4.2. Aqui basta observar o seguinte: *meus* pode ser precedido de *os*, como em *os meus vizinhos*; igualmente, podemos dizer *alguns meus vizinhos* e podemos, em certas circunstâncias, passar *meus* para depois de *vizinhos*, como em *vizinhos meus*, *alguns vizinhos meus* etc. Mas nada disso é possível com *um* no outro sintagma, como os exemplos abaixo deixam bem claro:

(44) * O um cachorro...
* Alguns uns cachorros...
* Cachorro um...
* Alguns cachorros uns...

(Relembro que o asterisco (*) é usado para marcar construções mal formadas.)

Como se vê, também *um* e *meus* nos exemplos vistos têm comportamento gramatical diferente; exprimimos isso na análise dizendo que têm **funções diferentes**.

Vou terminar esta seção resumindo o que se viu sobre a estrutura interna da oração: uma oração se estrutura em constituintes hierarquicamente organizados; e cada um desses constituintes possui um comportamento gramatical próprio, o que vale dizer que possui uma função própria.

3.2.2. Funções sintáticas na oração

Passemos agora às funções sintáticas propriamente ditas. Vamos estudar em primeiro lugar as funções sintáticas de nível oracional — isto é, as funções que podem ser desempenhadas pelos constituintes imediatos da oração. Esses constituintes são os que, numa árvore como a mostrada em (43), aparecem imediatamente abaixo do nódulo correspondente à oração. Assim, são constituintes da oração e de nenhuma outra unidade interior a ela; representam, por assim dizer, o primeiro corte realizado na estrutura oracional. Como se pode ver nos exemplos dados, "sujeito" e "objeto direto" são funções de constituintes de nível oracional.

3.2.2.1. Predicado e núcleo do predicado

Existem outras funções, evidentemente. Vamos começar pela função de **núcleo do predicado** (abreviadamente, **NdP**), que é desempenhada no exemplo (41) pelo verbo *arranjaram.*

O NdP é uma função que não será propriamente definida aqui. Acontece que é necessário estabelecer um ponto inicial que sirva de fundamento à cadeia de definições que vou propor para as diferentes funções. Vou então partir do seguinte postulado: o verbo desempenha na oração unicamente a função de núcleo do predicado; essa é a única função que um verbo pode desempenhar, e somente um verbo pode ser núcleo do predicado. Em outras palavras, o verbo é sempre o NdP da oração; e o NdP da oração é sempre um verbo.

Note-se que isso pressupõe que já se tenham condições de identificar uma palavra como pertencente à classe dos verbos. Essa classe é fácil de identificar, pois tem propriedades morfológicas muito singulares; dessa forma, a identificação do verbo não deve apresentar problemas de monta.

A aceitação desse postulado leva a análises diferentes da tradicional em alguns casos, como em

(45) Meu nariz está entupido.

onde *está* deve ser analisado como NdP. É bom que o leitor vá se acostumando desde já com a idéia de que a presente análise não tem compromisso de princípio com a da gramática tradicional. Chamar *está* em (45) de "núcleo" do predicado não quer dizer que seja essa a palavra mais importante, nem que esteja transmitindo a parte mais relevante da mensagem. Relembro (caso seja necessário) que estamos lidando, neste capítulo, com a sintaxe da oração — isto é, com sua organização formal, e não com sua interpretação semântica.

Vamos admitir, ainda, que o núcleo do predicado faz parte de um constituinte chamado **predicado** (**Pred**). Em frases como (45), então, o constituinte de nível oracional, estritamente falando, é o Pred, que contém o NdP (e nada mais).

Desse modo, na oração

(46) Batista detesta o dentista.

o predicado é *detesta*; e o núcleo do predicado é igualmente *detesta*, porque o predicado só tem um elemento. Essa dualidade de funções para um só elemento pode parecer gratuita no momento, mas veremos na seção seguinte que, em certos casos, o predicado é composto de mais de um elemento, sendo apenas um deles o NdP. Ficará clara, então, a necessidade de distinguir Pred de NdP.

3.2.2.2. Predicado complexo; auxiliar

Ao estabelecer a relação entre o NdP e o verbo, só dei exemplos com verbos *finitos*, isto é, em forma conjugável. Mas como se devem tratar as seqüências de verbos finitos + verbos em forma não-fmita (gerúndio, infinitivo, particípio)? Ou seja, na frase

(47) Sarita está dormindo.

há um ou dois núcleos do predicado? Temos aqui um problema a discutir; e, a bem dizer, esse problema não está cabalmente resolvido, no que pesem os diversos trabalhos que trataram dele nos últimos anos. Aqui vou apresentar uma espécie de compromisso, que espero seja adequado para a maioria dos casos, e que se harmoniza, no essencial, com a maioria das análises existentes.

Para começar, vou antecipar algumas informações a respeito do fenômeno da transitividade, a ser estudado no capítulo 6. O predicado faz certas exigências quanto aos complementos que o podem acompanhar na oração. Assim, há verbos, como *fazer*, que exigem a presença de um objeto direto; outros, como *falecer*, recusam a presença de objeto direto; e ainda há outros, como *comer*, que admitem o objeto direto opcionalmente, já que se pode dizer

(48) Sarita já comeu a empada.

e também

(49) Sarita já comeu.

Cada verbo tem, portanto, um conjunto de traços que especificam os complementos que ele exige, recusa ou aceita livremente; esse conjunto de traços exprime a transitividade do verbo (maiores detalhes no capítulo 6). Cada predicado (e, portanto, cada oração) possui um conjunto próprio de traços de transitividade, derivado da transitividade de seu verbo. Vendo o fato de outro ângulo, podemos dizer que em uma sentença há tantos conjuntos de traços de transitividade quantos predicados (ou orações) ela contém.

Mas em uma frase como

(47) Sarita está dormindo.

só se pode apurar **um** conjunto de traços de transitividade, aquele que se refere ao elemento *dormindo*. Isto é, a transitividade da seqüência *está dormindo* é idêntica, em todos os pontos, à do verbo *dormir* sem verbo auxiliar. Isso acontece sempre que ocorre uma seqüência de *estar* mais um verbo no gerúndio. A presença ou ausência de *estar* nessas estruturas não faz diferença para efeitos de aceitação ou recusa de complementos.

Ora, se quisermos manter o princípio de que cada predicado tem um conjunto próprio de traços de transitividade, deveremos admitir que há um só predicado — e, por conseguinte, um só NdP — na frase (47). Nesses casos, analisamos seqüências do tipo de *está dormindo* como predicados complexos, ou seja, compostos de NdP mais outro elemento ainda não definido.

Não se pode generalizar essa análise para qualquer seqüência de verbo finito + gerúndio. Em certos casos, a situação é diferente, obrigando-nos a aceitar a presença de dois NdPs e dois predicados separados. É o caso, por exemplo, de

(50) Toninho apanhou lutando.

À primeira vista, não haveria diferença entre (50) e (47): ambas apresentam um verbo em forma finita, seguido de um gerúndio. No entanto, no caso de (50), o verbo finito, *apanhou*, tem transitividade própria e pode receber complementos independentemente do outro verbo, *lutando*. Por exemplo, *lutar* não aceita complementação por um termo regido da preposição *de*, ao passo que *apanhar* aceita:

(51) * Toninho lutava de Sarita.
(52) Toninho apanhava de Sarita.

Ora, a frase (50) aceita o acréscimo de um termo com *de*:

(53) Toninho apanhou de Sarita lutando.

Como a possibilidade de ocorrência desse termo não pode ser atribuída ao verbo *lutar* [já que (51) é inaceitável], temos de admitir que o verbo

apanhar em (53) faz valer seus traços de transitividade. A diferença entre (47) e (50), portanto, é a seguinte: em (47), a transitividade de *está dormindo* é idêntica à de *dormia*, por exemplo; ou seja, *está* não influi em nada. Já em (50) cada verbo tem transitividade independente.

A conclusão é que, em (47), temos somente um predicado e, em (50), temos dois; e uma conseqüência é que em (47) há uma oração, e em (50) duas (quanto aos critérios de contagem de orações, ver a seção 5.2.1.).

Na frase

(47) Sarita está dormindo.

o núcleo do predicado é *dormindo*; ao elemento *está* atribuiremos a função de **auxiliar**.

Essa análise introduz uma complicação na conceituação de NdP dada na seção precedente. Eu disse ali que o NdP é o "verbo". Mas em (47) encontramos um verbo, *está*, que não é NdP. No entanto, há razões importantes para considerar *dormindo*, e não *está*, como NdP (essas razões serão vistas na seção 6.5.). Por outro lado, o número de verbos que podem funcionar como auxiliares é pequeno (ver lista abaixo), de modo que a complicação não é proibitiva. Vou, pois, completar a conceituação de NdP especificando que essa função é sempre desempenhada pelo verbo da oração, exceto nos casos em que há dois verbos em uma única oração. Nesses casos, o verbo em forma finita (conjugada) é o auxiliar, e o verbo em forma não-finita (particípio, gerúndio ou infinitivo) é o NdP.

O NdP de (47) é um gerúndio; mas é possível mostrar, com argumentação análoga, que há também NdPs com auxiliar formados por infinitivos ou por particípios, como em

(54) Sarita vai dormir.
(55) Sarita tem dormido.

Nesses casos, tal como em (47), o elemento conjugado (*vai, tem*) é irrelevante para efeitos de escolha de complementos e, portanto, não forma, por si só, um predicado. Os verbos que funcionam dessa maneira são denominados **verbos auxiliares**, ou simplesmente **auxiliares**. Em português, são poucos:

ir (+ infinitivo);
ter, haver (+ particípio);
estar, vir, ir, andar (+ gerúndio).

A esses é preciso acrescentar o verbo *ser* (mais raramente, *estar*) que ocorre na construção passiva:

ser, estar (+ particípio)

isso porque o verbo *ser* nas passivas também não apresenta transitividade própria, e além do mais a seqüência *ser* + particípio parece funcionar como um constituinte.

E, finalmente, há ainda um grupo de auxiliares (tradicionalmente chamados "modais" e "aspectuais") que se constroem todos

com infinitivo, e alguns com *de*, *que* ou *a*:

poder
dever
acabar de
deixar de
começar a
continuar a
ter de/que
haver de/que

É necessário considerá-los também auxiliares porque, quando seguidos de infinitivo, não apresentam traços próprios de transitividade.

É bom lembrar que todos esses verbos podem igualmente aparecer em construções nas quais não são auxiliares; portanto, a identificação de um auxiliar (e de um predicado complexo) exige sempre algum cuidado.

Como se vê, cada tipo de complemento coocorre obrigatoriamente com um de seus auxiliares específicos. Isto é, sempre que o auxiliar é o verbo *ir*, o NdP deve estar no infinitivo:

(56) Manuel *vai cortar* o bigode.

O auxiliar sendo *ter*, o NdP estará no particípio:

(57) Manuel *tem falado* mal de você.

E se o auxiliar for *estar*, *andar*, o NdP deverá estar no gerúndio:

(58) Manuel está penteando o bigode.

O predicado complexo é, pois, sempre composto de auxiliar (Aux) mais NdP; e pode haver mais de um auxiliar, posicionados sempre segundo uma ordenação rígida (que é a mesma mostrada na lista acima: primeiro o auxiliar seguido de infinitivo, depois o auxiliar seguido de particípio e, por último, o auxiliar seguido de gerúndio). Se quisermos fazer uma frase utilizando vários auxiliares, teremos de atender às duas exigências acima formuladas: primeiro, a coocorrência de cada tipo de Aux com um de seus verbos auxiliares específicos; e, depois, a ordenação *infinitivo — particípio — gerúndio*.

Isso se consegue colocando o próprio auxiliar no infinitivo ou no particípio quando necessário. Digamos, então, que vamos montar uma frase com um Aux no infinitivo e outro no gerúndio; o resultado deverá ser algo como

(59) Manuel vai estar contando piadas.

Essa frase é bem formada porque: (a) o Aux no infinitivo (*estar*) é precedido de uma forma de *ir*, e o Aux no gerúndio é precedido de uma forma de *estar*; e (b) o infinitivo vem antes do gerúndio. Qualquer desobediência a essas regras dá como resultado uma frase mal formada ou, então, uma frase onde os dois verbos não formam um predicado complexo. Por exemplo, desobedecendo às condições de coocorrên-

cia, poderemos criar a estrutura mal formada

(60) * Manuel tem contando piadas.

E desobedecendo à condição de ordenação poderemos criar a estrutura (igualmente mal formada)

(61) * Manuel está tendo contado piadas.

3.2.2.3. Sujeito

Agora temos um ponto de partida para investigar a organização da estrutura oracional. Verificaremos que cada um dos constituintes em que se divide a oração tem um comportamento gramatical próprio, ou seja, cada um desempenha uma função própria. Há também alguns casos de dois ou mais constituintes de comportamento semelhante; nesses casos, alguma função poderá ser repetida. Isso não acontece com qualquer função, porém; só algumas funções podem ocorrer mais de uma vez em uma estrutura, e isso deverá ser apurado e explicitado ao se definir cada uma das funções.

Nossa primeira observação será a seguinte: conforme já foi brevemente mencionado acima, existe, na maioria das orações, um constituinte que se harmoniza com o NdP em número e pessoa. Vejamos primeiramente o que vem a ser esse fenômeno de harmonização em número e pessoa.

Seja um exemplo como o seguinte:

(62) Meus sobrinhos comeram a melancia.

É fácil verificar que a forma do verbo (do NdP), *comeram*, depende, de certa forma, de traços do constituinte *meus sobrinhos*. Assim, se no lugar de *meus sobrinhos* colocarmos o singular, *meu sobrinho*, o verbo terá de se adaptar:

(63) Meu sobrinho comeu a melancia.

Por outro lado, uma mudança semelhante no constituinte *a melancia* não afeta a forma do verbo:

(64) a. Meus sobrinhos comeram a melancia.
b. Meus sobrinhos comeram as melancias.
c. Meu sobrinho comeu a melancia.
d. Meu sobrinho comeu as melancias.

Ou seja, existe uma espécie de harmonia formal entre o verbo e o constituinte *meu(s) sobrinho(s)*, mas não há nada parecido entre o verbo e *a(s) melancia(s)*. Dizemos que *meu(s) sobrinho(s)* e o verbo estão em **relação de concordância** (ou, simplesmente, que **concordam**).

Conforme foi adiantado, a concordância não funciona apenas para o **número** (singular, plural), mas também vale para a pessoa. A noção de "pessoa" será estudada em

outro local (seção 6.6.), quando examinarmos mais cuidadosamente o mecanismo da concordância. Aqui, basta saber que a maioria dos sintagmas nominais são considerados como de "terceira pessoa"; e algumas formas especiais são de "primeira pessoa" (*eu, nós*), ou de "segunda pessoa" (*tu, vós*). A segunda pessoa é de uso muito raro no português padrão do Brasil, e por isso a exemplificação utilizará apenas a primeira e a terceira.

O sintagma *meus sobrinhos* é de terceira pessoa (assim como a imensa maioria dos sintagmas da língua). Isso significa, na verdade, que, quando o verbo concorda com ele, deve assumir uma dentre várias formas igualmente denominadas "de terceira pessoa"; exemplos são (no plural): *comeram, dormem, serão*:

(65) Meus sobrinhos dormem no segundo andar.
(66) Meus sobrinhos serão grandes tenistas.

Se substituirmos *meus sobrinhos* por uma das poucas formas rotuladas como de primeira pessoa, o verbo deverá assumir formas diferentes; existe, pois, um outro conjunto de formas verbais, denominadas "de primeira pessoa". Por exemplo,

(67) Nós comemos a melancia.
(68) Nós dormimos no segundo andar.
(69) Nós seremos grandes tenistas.

As formas *comemos, dormimos, seremos* são formas de primeira pessoa.

Voltando às funções sintáticas, então, verificamos que na oração

(62) Meus sobrinhos comeram a melancia.

os sintagmas *meus sobrinhos* e *a melancia* têm comportamento gramatical diferente: *meus sobrinhos* está em relação de concordância com o NdP, e *a melancia* não está — portanto, cada um desses sintagmas tem uma função diferente. A função desempenhada nessa frase por *meus sobrinhos* é denominada **sujeito** (abreviadamente, Suj): dizemos que *meus sobrinhos* é o sujeito da oração (62).

Vou, então, definir essa função da seguinte maneira:

Sujeito é o termo da oração que está em relação de concordância com o NdP.

Essa será nossa definição de sujeito. É uma definição formal e não diz nada a respeito do papel semântico ou discursivo do termo em questão; em outras palavras, não estamos aqui preocupados com o termo que exprime o agente de uma ação, nem com o termo que exprime a entidade sobre a qual se faz uma declaração. Trata-se simplesmente de um dos constituintes da oração, vinculado a ela através de uma relação formal bem definida. A função de sujeito é um dos aspectos da organização formal da oração, e

não um dos aspectos da mensagem veiculada pela oração. No capítulo 10, veremos qual a relação que existe entre o sujeito de uma oração e noções semânticas, tais como "aquele que pratica a ação". Essa noção é também importante, mas deve ser colocada em seu componente próprio da descrição gramatical.

A propriedade "estar em relação de concordância com o NdP", que define o sujeito, é também chamada um **traço** que o constituinte tem na oração. Esse traço é abreviado assim: [+CV] (CV significa "concordância verbal"); dizemos, então, que o sintagma *meus sobrinhos* em (62) é marcado com o traço [+CV]; os outros sintagmas, que não estão em relação de concordância com o NdP, são todos marcados [-CV]. Essa notação por traços apresenta vantagens, de maneira que vai ser utilizada com freqüência neste livro. A análise sintática será, na verdade, expressa principalmente em termos de traços atribuídos a determinados constituintes; e cada traço exprime um aspecto do comportamento sintático do constituinte em questão.

Observe-se que a definição adotada de "sujeito" nos obriga a analisar como sem sujeito frases do tipo

(70) Vendi meu jegue.

A solução tradicional é considerar que há aí um "sujeito oculto"; no entanto, essa análise é inconsistente com a definição de sujeito como termo que está em relação de concordância com o verbo. A justificação detalhada da análise de (70) como oração sem sujeito envolve certas noções semânticas, motivo pelo qual será adiada para o próximo capítulo. No Quadro 10-C, a questão é retomada e discutida; por ora, lembremo-nos que (70) é uma oração sem sujeito, já que não existe aí nenhum termo explícito que esteja em relação de concordância com o verbo.

A aplicação da definição de sujeito, a bem dizer, deixa certos casos duvidosos. Talvez o mais sério seja o do gerúndio, que não comporta desinências de pessoa-número, mas ainda assim é usualmente analisado como tendo sujeito em frases como

(71) Marivânia chegando, a farra vai começar.

Esses casos foram discutidos em um livro anterior (Perini, 1989); vou dar aqui um sumário das conclusões, com alguma argumentação.

A rigor, seria necessário negar que haja sujeito na primeira oração em (71), já que aí o verbo não concorda com nenhum dos termos. No entanto, sob vários outros pontos de vista, *Marivânia* funciona como se fosse um sujeito, o que cria um problema quanto a sua análise.

Acontece que o sujeito, definido por sua relação de concordância

com o verbo, apresenta também alguns outros traços que, se não estão presentes em todos os casos, estão presentes na maioria deles; por conseguinte, esses traços contribuem para delinear o protótipo da função que chamamos **sujeito**. O mais óbvio desses traços é a posição logo antes do NdP: sabemos que essa é a posição mais natural do sujeito na maioria das frases. Em (71), o termo *Marivânia* está nessa posição, justificando pelo menos a suspeita de que se trate de um sujeito.

Mas o mais significativo é que as condições que governam a possibilidade de ocorrência do sujeito antes ou depois do NdP se aplicam igualmente aos casos evidentes de sujeito e a casos como o de *Marivânia* em (71). Assim, há casos em que não se permite a posposição do sujeito; podemos dizer

(72) Chegou um amigo meu de Cuiabá.

mas não

(73) * Está desenhando um sobrinho seu na biblioteca.

Ora, os fatos são exatamente paralelos com o SN que acompanha um gerúndio:

(74) Chegando um amigo meu, por favor receba-o bem.

(75) * Desenhando um sobrinho meu na biblioteca, tive que ficar no quarto.

Isso certamente sugere que o gerúndio tem sujeito. Ou seja, se o gerúndio tiver sujeito, ficará mais simples a descrição das condições que governam a possibilidade de posposição de elementos, como *um amigo meu* e *um sobrinho meu* nas frases (74) e (75). Já se analisássemos esses termos como desempenhando outra função qualquer, teríamos de considerar esse um estranho caso de coincidência entre o comportamento do sujeito e o dessa outra função.

Outra razão é que o sujeito, como sabemos, pode ser retomado por pronome em caso reto: *eu*, por exemplo, e não *me* ou *mim*. Desse ponto de vista, também o SN que acompanha o gerúndio parece ser um sujeito:

(76) Eu chegando, a farra vai começar.

Por essas razões, sugiro que *Marivânia* em (71), assim como *eu* em (76), seja analisado como sujeito. Isso naturalmente complicará a conceituação de sujeito, já que não se poderá dizer simplesmente que se trata do termo que está em relação de concordância com o verbo; mas traz vantagens, porque nos permite analisar as condições de posposição de sujeito e de ocorrência dos pronomes retos de maneira mais simples e unificada.

Quanto às razões para a inexistência de concordância no caso dos gerúndios, ver Perini, 1989, p. 95-96.

3.2.2.4. Objeto direto

Voltemos ao exemplo

(62) Meus sobrinhos comeram a melancia.

Vamos agora ver que função deve ser atribuída ao sintagma *a melancia*. Já sabemos que não é sujeito, porque é marcado (nessa frase) como [–CV] — ou seja, não está em relação de concordância com o verbo.

A gramática tradicional distingue várias funções dentre as que marcaremos como [–CV]. Assim, temos casos como

(62) Meus sobrinhos comeram *a melancia.*

(77) Todos acharam esse livro *uma droga.*

(78) Jeremias reclama *freqüentemente.*

Tradicionalmente, analisa-se o elemento grifado de (62) como "objeto direto"; o de (77) como "predicativo do objeto"; e o de (78) como "adjunto adverbial". Existem intuições provavelmente corretas por trás dessa análise; vou tentar encontrar justificativas formais para ela.

Vamos começar com o caso de *a melancia* em (62), que seria um "objeto direto". Procuraremos propriedades sintáticas (traços) desse constituinte, as quais poderão servir de base para definir uma nova função.

Podemos começar observando o seguinte: existe uma oração **correspondente** a (62), na qual o constituinte *a melancia* se encontra deslocado para o início da oração:

(79) A melancia, meus sobrinhos comeram.

(A noção de **correspondência** foi introduzida na seção 2.1.)

Esse é um traço do constituinte *a melancia* em (62); podemos dizer que esse elemento é marcado [+Ant] ("Ant" quer dizer "anteposição"). Há elementos da oração que não podem aparecer antepostos em frases correspondentes e que, portanto, são marcados [–Ant]:

(80) * Comeram, meus sobrinhos a melancia.

(81) * Uma droga, todos acharam esse livro.

Vamos admitir que esse traço é característico do objeto direto (abreviadamente, OD), que se definirá, por enquanto, através do conjunto de traços [–CV, +Ant]. O objeto direto se distingue do sujeito porque este é [+CV]; e se distingue da função representada por *uma droga* em (81) porque esta é [–Ant].

Por ora, conseguimos distinguir o "objeto direto" tradicional do "predicativo do objeto", através do traço [Ant]. Mas ainda não fizemos a distinção entre o objeto direto e o "adjunto adverbial" que estaria presente em

(78) Jeremias reclama freqüentemente.

Como decidimos dar um crédito à análise tradicional, vamos procurar um meio de distinguir a função de *freqüentemente* em (78) da função de *a melancia* em (62).

O traço que realiza essa distinção se baseia na seguinte observação: certos constituintes da oração podem ser retomados em perguntas por meio do elemento *que* (ou *o que*, ou ainda *quem*). Ou seja, pode-se definir uma relação discursiva tal que uma frase possa ser resposta adequada a uma pergunta que é introduzida por um desses elementos interrogativos. As duas (pergunta e resposta) se relacionam formalmente porque na pergunta o elemento *que* substitui o constituinte em exame (com ou sem inversão de ordem). Assim, temos o par

(82) a. — O que meus sobrinhos comeram?
b. — Meus sobrinhos comeram a melancia.

Creio que qualquer pessoa perceberá que (82b) é uma resposta adequada à pergunta (82a). Agora veja-se o seguinte par:

(83) a. — O que Jeremias reclama?
b. — Jeremias reclama freqüentemente.

Acontece que (83b) não é uma resposta adequada para (83a). Dizemos, nesses casos, que *freqüentemente* não pode ser retomado pelo elemento *o que*; o contrário acontece com *a melancia* em (62), que pode ser retomado pelo elemento *o que*.

Esse traço é que distingue a função de *a melancia* em (62), que é marcada [+Q], da função de *freqüentemente* em (78), que é marcada [–Q].

Antes de passar adiante, vou abordar rapidamente alguns pontos de interesse. Em primeiro lugar, é necessário decidir a função que deve ser atribuída ao constituinte anteposto *a melancia* em

(79) A melancia, meus sobrinhos comeram.

Sabemos que na oração correspondente (62), em que *a melancia* aparece no final, a função é de objeto direto; pergunta-se agora se deve continuar sendo objeto direto em (79), ou se a diferença de posição deve valer para que se defina uma nova função.

A resposta será que *a melancia* é objeto direto tanto em (62) quanto em (79); ou seja, a anteposição não afeta a função do objeto direto. As razões para se adotar essa solução estão explicitadas no Quadro 7-B. Essa conclusão vale, obviamente, para todas as demais funções que podem aparecer deslocadas de seus lugares em frases correspondentes.

3.2.2.5. Complemento do predicado

Utilizando os traços vistos até agora, o sintagma *um artista* em

(84) Vincent é um artista.

deveria ser analisado como objeto direto, pois, como é fácil verificar, é marcado com os traços [–CV, +Ant, +Q]. Essa análise contraria a da gramática tradicional, que analisa tais casos como de "predicativo do sujeito". Na verdade, como já apontei em outro livro (Perini, 1989), não é muito fácil distinguir essas duas funções. Na análise anterior, optei por não as separar, de maneira que *um artista* em (84) foi realmente analisado como objeto direto. Agora, no entanto, depois de um exame mais cuidadoso da questão, decidi alterar minha posição, mantendo a distinção tradicional e definindo-a através de um novo traço. Como se verá, permanecem alguns problemas; mas agora me parece que o mais indicado é manter a distinção entre o objeto direto e a função desempenhada por *um artista* em (84).

O novo traço se baseia na relação de concordância nominal que se observa com freqüência entre o sujeito e o complemento de verbos como *ser*, que vou denominar **complemento do predicado (CP)** (evitando o uso do termo "predicativo", já utilizado para designar outra função; ver a seção 3.2.2.6.). Observa-se que, quando o CP é representado por um item passível de concordância nominal, ele concorda com o sujeito:

(85) Carolina está entusiasmada.

(86) Os jogadores estão entusiasmados.

Por isso, diremos que os termos *entusiasmada, entusiasmados* são marcados [+CN]. Isso estabelece a distinção entre CP e o objeto direto, que não mostra concordância nem mesmo nos raros casos em que é representado por item apropriado:

(87) Carolina sonhou colorido.

Dessa maneira, temos aqui um novo traço distintivo, [CN], e uma nova função, o complemento do predicado, cuja definição até o momento é [–CV, +Ant, +Q, +CN]. O objeto direto se define como [–CV, +Ant, +Q, –CN]. Estou definindo o traço [CN] em função da possibilidade de concordância nominal com algum outro termo da oração, não necessariamente o sujeito; isso porque, como veremos em 3.2.2.6., deve-se atribuir o traço [+CN] a certos termos que concordam com o objeto direto.

A identificação do CP apresenta determinado número de problemas práticos. O primeiro deles diz respeito a limitações morfológicas de muitos itens léxicos. Assim, em

(84) Vincent é um artista.

não se pode dizer que haja concordância nominal explícita entre o sujeito e o CP; no caso, ambos são masculinos e singulares, mas isso não é necessário:

(88) Vincent é a próxima testemunha.

Nesses casos, fica difícil demonstrar diretamente que *a próxima testemunha* é CP. É importante notar que não há tampouco discordância, porque *a próxima testemunha* é um SN e tem, por isso, gênero e número próprios, não atribuídos por concordância. O gênero e o número são determinados, para os SNs, por razões semânticas, quando cabe; ou por marca idiossincrática, como em (88): *testemunha* é idiossincraticamente marcado como feminino.

A saída é lançar mão da transitividade normal do verbo como pista. Esse procedimento é relativamente fácil porque há poucos verbos que aceitam CP ou OD, indiferentemente. Um desses verbos é possivelmente *virar*, como em

(89) Ela virou a picanha em cima da grelha. (OD)
(90) Ela virou um monstro depois da plástica. (CP)

Mas não há realmente certeza de que se trate de um só verbo; talvez seja um caso de homonímia. Por ora, portanto, deixaremos de lado esses casos problemáticos.

Os verbos que admitem CP são relativamente poucos; podemos citar

ser, estar, parecer, continuar, ficar, virar, permanecer, chamar-se, tornar-se, sentir-se.

Dessa maneira, é preciso reconhecer que a identificação do CP apresenta dificuldades; por outro lado, é importante distinguir o CP do OD porque, nos casos claros, existe uma diferença de comportamento desses dois termos, e essa diferença se liga ao importante fenômeno da concordância nominal.

Em uma frase como

(84) Vincent é um artista.

não se pode observar diretamente a concordância nominal entre o sujeito e o CP. Ainda assim, diremos que *um artista* é [+CN]; a concordância não se manifesta aí por causa das limitações morfológicas da palavra *artista*, que não admite variação derivada (isto é, *artista* tem gênero e número, mas estes são propriedades inerentes à palavra, e não resultado de concordância). Uma solução semelhante foi proposta na seção 3.2.2.3., para podermos aceitar que os gerúndios têm sujeito, embora, por razões morfológicas, não manifestem a concordância verbal. Reconhecendo embora a inconveniência ocasional dessa solução, vou adotá-la como a mais indicada no caso.

Um segundo problema ligado à identificação do CP provém de frases como

(91) Cerveja é bom para lavar o cabelo.

Essas construções são de funcionamento misterioso, pois o complemento é passível de concordância e, no entanto, esta é bloqueada:

(92) ?? Cerveja é boa para lavar o cabelo.

Isso parece acontecer somente quando o sujeito é representado por um núcleo (NSN) sem nenhum termo acompanhante:

(93) Essa cerveja é boa para lavar o cabelo.
(94) * Essa cerveja é bom para lavar o cabelo.

Não parece tratar-se de alguma restrição semântica, porque não se pode atribuir importe semântico ao gênero: que significado se poderia atribuir à diferença de gênero entre *violão* e *guitarra*? Este problema deverá ficar de lado, à espera de estudos que o esclareçam.

3.2.2.6. Atributo e predicativo

Há ainda duas outras funções que se podem considerar definidas, com nossos quatro traços. Vimos que o sujeito se define como [+CV]; que o objeto direto se define como [–CV, +Ant, +Q, –CN]; e que o complemento do predicado se define como [–CV, +Ant, +Q, +CN]. Agora: nas orações (77) e (78) temos duas outras funções que já podem ser formalmente definidas. Vejamos o exemplo

(77) Todos acharam esse livro uma droga.

O constituinte *uma droga* é marcado [–CV], pois não está em relação de concordância com o NdP; é marcado [–Ant], porque não pode aparecer anteposto em uma frase correspondente:

(81) * Uma droga, todos acharam esse livro.

É marcado [+Q] porque pode ser retomado por *o que*:

(95) a. — Todos acharam esse livro o quê?
b. — Todos acharam esse livro uma droga.

(95b) é uma resposta adequada a (95a).

E é marcado [+CN] porque, quando representado por um item passível de concordância nominal, esta se manifesta (com o objeto direto):

(96) Todos acharam esse livro péssimo.
(97) Todos acharam essa peça péssima.

Temos, portanto, uma nova função, definida pelos traços [–CV, –Ant, +Q, +CN]. Vou chamá-la **predicativo** (Pv).

Passemos agora ao exemplo

(78) Jeremias reclama freqüentemente.

O constituinte *freqüentemente* é marcado [–CV], por razões óbvias; é [+Ant], porque (98) é aceitável:

(98) Freqüentemente, Jeremias reclama.

E é marcado [–Q] porque, como já vimos, (86b) não é uma resposta adequada a (86a):

(83) a. — O que Jeremias reclama?
b. — Jeremias reclama freqüentemente.

A função desempenhada por *freqüentemente* em (78) se chama **atributo** (Atr) e se define pelos traços [–CV, +Ant, –Q].

E quanto ao traço [CN]? Já vimos que esse traço só pode ser observado em certos casos. Se substituirmos *freqüentemente* por *indignado*, verificaremos que se trata ainda de um atributo, pois deve ser marcado [–CV, +Ant, –Q]. E, além disso, é evidentemente [+CN]; vamos, pois, acrescentar mais esse traço à definição de atributo: [–CV, +Ant, –Q, +CN].

3.2.2.7. Negação verbal

As cinco funções estudadas até o momento (sujeito, objeto direto, complemento do predicado, predicativo e atributo) são desempenhadas cada uma por um número muito grande de sintagmas; não haveria meios de elaborar, por exemplo, a lista de **todos** os possíveis sujeitos do português. Sabemos que a lista não acabaria nunca, pois incluiria não apenas palavras individuais, mas um número ilimitado de sintagmas maiores.

Isso é o que acontece com muitas funções sintáticas. Por outro lado, existem outras funções que só podem ser desempenhadas por um número limitado de elementos, de modo que é possível dar a lista de todos eles. É o que acontece com a função que veremos agora, denominada **negação verbal** (NV). Essa função é exemplificada pela palavra *não* na frase

(99) O ministro não aprecia jiló.

Essa função, como é fácil verificar, é marcada negativamente para os traços já vistos: [–CV, –Ant, –Q]. Com isso, a negação verbal fica diferenciada das demais funções estudadas. No entanto, essa função tem uma característica mais importante, que merece ser expressa em um traço especial: a negação verbal só pode ocorrer **logo antes do NdP**, sem possibilidade de inserção de nenhum elemento entre os dois. Vou abreviar essa propriedade através da notação [pNdP]; assim, a negação verbal é marcada [+pNdP]; todas as outras funções vistas, claro, serão [–pNdP]. É obviamente impossível a existência, em uma mesma oração, de dois constituintes marcados [+pNdP], pois **ambos** teriam de ficar imediatamente antes do NdP.

Entre a negação verbal e o NdP não podem ocorrer nem mesmo elementos parentéticos separados por vírgula, como os dos exemplos

(100) O ministro, dizem, não aprecia jiló.
(101) O ministro não aprecia, dizem, jiló.

Mas:

(102) * O ministro não, dizem, aprecia jiló.

A bem dizer, existe um pequeno grupo de elementos que podem aparecer entre a negação verbal e o NdP, a saber, os pronomes clíticos como *me, nos, lhe* etc.:

(103) O ministro não nos recebeu na hora marcada.

Mas, como se trata de um grupo bem delimitado, o fato não prejudica nossa definição; basta acrescentar essa ressalva à formulação do traço. O comportamento dos pronomes clíticos, como se sabe, é muito peculiar; será estudado no capítulo 8.

Aparentemente, existem apenas duas palavras em português que podem desempenhar a função de negação verbal: *não* e *mal*, em uma de suas acepções:

(104) A noiva mal chegou a tempo para o casório.

Há um grupo de itens que, provavelmente por razões de semelhança semântica, costumam ser classificados juntamente com *não* e *mal*; no entanto, seu comportamento sintático é claramente diferente do de *não* e *mal*. Exemplos são: *nunca, jamais, já, nada, ninguém.*

Uma última observação sobre a negação verbal: o fato de que ela não se pode separar do NdP pela inserção de nenhum elemento (exceto clíticos) sugere que ela seja também parte do predicado — não é, estritamente falando, um constituinte de nível oracional.

3.2.2.8. Adjunto adverbial, adjunto oracional, adjunto circunstancial

A gramática tradicional engloba, sob o rótulo "adjunto adverbial", um conjunto bastante variado de funções. Já vimos que uma dessas é a que chamamos aqui "atributo"; agora veremos outras três funções, igualmente denominadas "ad-juntos adverbiais" pela análise tradicional.

Vejamos em primeiro lugar o comportamento do constituinte *completamente* em

(105) Miguel decorou o apartamento completamente.

Esse elemento tem os traços [–CV, –Q], tal como o atributo; mas, ao contrário dele, é [–Ant]:

(106) * Completamente, Miguel decorou o apartamento.

Trata-se, portanto, de uma nova função, caracterizada, entre outras coisas, por uma posição relativamente fixa na oração. Chamo-a **adjunto adverbial** (AA); define-se pelos traços [–CV, –Ant, –Q, –CN, –NdP]. Relembro que, apesar da identidade de nomes, o adjunto adverbial só compreende uma pe-

quena parte dos termos tradicionalmente chamados "adjuntos adverbiais".

Passemos agora ao caso de *francamente* em

(107) Esse professor, francamente, é um neurótico.

Esse termo tem alguma semelhança com o atributo, porque é marcado [–CV, +Ant, –Q]; mas é aparentemente marcado [–CN], o que o diferencia do atributo. Além disso, *francamente* em (107) não parece compor constituinte com nenhum outro elemento, ao passo que em um caso claro de atributo, como

(78) Jeremias reclama freqüentemente.

freqüentemente faz constituinte com *reclama*, assim:

(108) [Jeremias] — [reclama freqüentemente]

Na literatura lingüística moderna, elementos como *francamente* são considerados à parte, como um elemento anexo à oração, talvez mesmo externo a ela; denominam-se "advérbios de oração" (*sentence adverbs*, em inglês). Aqui, a função correspondente será chamada **adjunto oracional** (AO).

Existe um traço adicional (além de [CN]) que distingue o adjunto oracional do atributo. Para defini-lo, vamos considerar brevemente o fenômeno da "clivagem". Trata-se de uma construção que põe em evidência um elemento da oração, com o auxílio do verbo *ser* mais o item *que*; pode-se, assim, "clivar" um sujeito, como em

(109) Foram meus sobrinhos que comeram a melancia.

ou então um objeto, como em

(110) Foi a melancia que meus sobrinhos comeram.

e assim por diante.

Digamos, então, que os constituintes que podem ser clivados como nos exemplos acima são marcados com o traço [+Cl]; constituintes não-cliváveis são marcados [–Cl].

Ora, das funções vistas até o momento, todas podem ser clivadas, exceto apenas as de NdP e de negação verbal (que já são em si suficientemente idiossincráticas). Em particular, o atributo pode ser clivado:

(111) É o dia inteiro que Jeremias reclama.

Mas o adjunto oracional é claramente não-clivável, como se vê em

(112) * É francamente que esse professor é um neurótico.

Dessa maneira, vamos utilizar o traço [Cl] para caracterizar o adjunto oracional. O adjunto oracional, portanto, se define como [–CV, +Ant, –Q, –CN, –Cl, –pNdP].

Examinemos agora o comportamento de *muito* em

(113) Juracy bebe muito.

Até onde podemos verificar, seria um atributo, pois é marcado com os traços [–CV, +Ant, –Q, +Cl, –pNdP], e o traço [CN] pode não estar explicitado por causa da morfologia de *muito*, que é invariável. Teria então a mesma função que *freqüentemente* em

(78) Jeremias reclama freqüentemente.

Há, porém, uma diferença entre os dois casos no que se refere a suas possibilidades de ocorrência em várias posições na oração. Sabemos que ambos podem ser antepostos (levados para o início da oração) e, por isso, são marcados [+Ant]. Mas só o segundo pode ser colocado imediatamente antes do NdP (ou da negação verbal, se houver):

(114) Jeremias freqüentemente reclama.

(115) * Juracy muito bebe.

Temos aqui base para a postulação de um novo traço, a que chamarei [PA] (de "posição do auxiliar", denominação devida a Jackendoff, 1972, p. 49). O traço [PA] exprime a propriedade de ocorrer entre o sujeito e o NdP (ou entre o sujeito e a negação verbal mais o NdP, se for o caso).

Além disso, voltando ao traço [CN], observa-se que, quando a concordância nominal é possível, o transporte para a posição do auxiliar também é possível:

(116) Jeremias reclamou indignado.

(117) Jeremias indignado reclamou.

Isso sustenta a idéia de que temos um atributo em (78), mas não em (113).

Temos, assim, condições de diferenciar a função de *muito* em (113) da de *freqüentemente* em (78): no primeiro caso, temos uma função definida como [–CV, +Ant, –Q, +Cl, –CN, –pNdP, –PA]; vamos chamar a essa função **adjunto circunstancial** (AC). No segundo caso, a função é definida como [–CV, +Ant, –Q, +Cl, +CN, –pNdP, +PA]; e trata-se, como sabemos, de um atributo.

Devo avisar desde já que, dos traços utilizados nesta análise, [PA] é provavelmente o menos satisfatório; e isso principalmente porque são muito numerosos os casos em que há variação ou insegurança no julgamento dos falantes. Por exemplo, na oração

(118) Jeremias reclama o dia inteiro.

não é claro para todos os falantes se *o dia inteiro* pode ou não pode ocorrer na posição de auxiliar. O leitor poderá verificar isso por si mesmo, pedindo a algumas pessoas que dêem seu julgamento de aceitabilidade sobre a frase

(119) ? Jeremias o dia inteiro reclama.

De qualquer maneira, como a presente análise não tem a pretensão de ser perfeita, ficaremos com o traço [PA], no que pesem suas deficiências, como parte da distinção entre o atributo e o adjunto circunstancial.

3.2.2.9. Funções repetidas na oração

Uma questão que não está satisfatoriamente investigada é a da possibilidade de se encontrar mais de uma ocorrência da mesma função em uma oração. Por um lado, é bastante claro que uma oração nunca pode possuir mais de um sujeito, objeto direto, complemento do predicado, predicado ou predicativo. Mas, para as outras funções, a possibilidade de repetição ainda não foi esclarecida.

Pelo menos uma função pode certamente aparecer repetida: o adjunto circunstancial (AC). Assim, podemos ter um exemplo como

(120) Ele se deixou levar ao desespero pelos credores.

Seguindo nossa análise, temos aqui dois ACs, *ao desespero* e *pelos credores.* Desse modo, podemos deixar estabelecido que o AC pode ocorrer mais de uma vez na mesma oração.

Mas é bom notar que justamente o AC é suspeito de ser um rótulo escondendo mais de uma função. Sabemos que a análise tradicional distinguiria *pelos credores* ("agente da passiva") de *ao desespero* (provavelmente "adjunto adverbial"); e pode haver alguma verdade nisso.

Quanto ao atributo, ao adjunto adverbial e ao adjunto oracional, faz falta um trabalho que deixe claro se podem ou não ocorrer repetidos na mesma oração. Aqui só nos será possível deixar a pergunta formulada; é uma das muitas dúvidas a respeito da estrutura da oração que permanecem, por ora, sem resposta.

3.2.2.10. Sumário: a estrutura da oração

A esta altura já podemos fazer um sumário geral da análise da oração. Não acredito que haja realmente esgotado a lista das funções de nível oracional. Mas as funções definidas nas seções acima bastam para fundamentar uma análise mais adequada e coerente da estrutura da oração do que a análise oferecida pela gramática tradicional.

Vimos que a oração simples em português se compõe de um conjunto de constituintes de comportamento sintático variado. Esse comportamento sintático pode ser descrito através de sete traços distintivos, listados e definidos na tabela a seguir.

Tabela 1: Traços distintivos

[CV] — A propriedade de estar em relação de concordância com o NdP.

[Ant] — A propriedade de poder aparecer no início da oração em uma frase correspondente.

[Q] — A propriedade de poder ser retomado pelos elementos *que*, *o que* ou *quem*.

[CN] — A propriedade de estar em relação de concordância (nominal) com outro termo da oração.

[Cl] — A propriedade de poder ocorrer como foco de uma frase clivada correspondente.

[PA] — A propriedade de poder ocorrer na posição do auxiliar (entre o sujeito e o NdP).

[pNdP] — A propriedade de *só* poder ocorrer imediatamente antes do NdP.

NOTA

Relembro que um constituinte é marcado positivamente(+) quando tem a propriedade descrita pelo traço, e negativamente (–) quando não a tem.

E definimos nove funções (além das de NdP, Aux e Pred), a saber:

Tabela 2: Funções de nível oracional

Sujeito, que se define como [+CV].
Objeto direto, ... [–CV, +Ant, +Q, –CN, +Cl, –PA].
Complemento do predicado, ... [–CV, +Ant, +Q, +CN, +Cl, –PA].
Predicativo, ... [–CV, –Ant, +Q, +CN, +Cl, –PA].
Atributo, ... [–CV, +Ant, –Q, +CN, +Cl, +PA].
Negação verbal, ... [+pNdP ...].
Adjunto adverbial, ... [–CV, -Ant, –Q, –CN, +Cl, –PA].
Adjunto oracional, ... [–CV, +Ant, –Q, –CN, –Cl, +PA].
Adjunto circunstancial, ... [–CV, +Ant, –Q, –CN, +Cl, –PA].
Núcleo do predicado (postulado).

NOTAS

(a) O sujeito é também [+Ant, +Q, +Cl, –pNdP], mas esses traços não são cruciais para sua identificação, porque o sujeito é a única função marcada [+CV].

(b) O NdP, o auxiliar e a negação verbal possivelmente não são funções de nível oracional, mas fazem parte do **predicado**.

3.3. VOCATIVO

Finalmente, vejamos o caso do vocativo, um termo geralmente considerado estranho à estrutura da oração. Como veremos, essa análise tradicional é correta.

À primeira vista, o vocativo parece merecer o rótulo de adjunto oracional:

(121) Serginho, a bandeira está no chão.

É fácil ver que o constituinte *Serginho* tem os traços [−CV, +Ant, −Q, −Cl, +PA], o que faria dele um AO.

No entanto, há certas considerações que nos levarão não só a negar que o vocativo seja um caso de AO, mas ainda a afirmar que a análise por traços sintáticos não se aplica a ele, porque sua conexão com a oração não é propriamente sintática. Quero dizer que a ligação entre o vocativo e a oração junto à qual ele pode ocorrer não tem a ver com a estrutura da própria oração, mas com a organização do discurso.

Há várias indicações formais que sugerem fortemente que esse é o caso. Em primeiro lugar, o vocativo pode separar-se da oração não apenas por vírgula, como está em (121), mas também por sinalização de final de período:

(122) Serginho! A bandeira está no chão.

Na fala, isso se traduz pela possibilidade de uma pausa de duração indefinida entre os dois elementos.

Depois, o vocativo pode estar separado da oração por uma mudança de interlocutor, sem que isso produza impressão nítida de interrupção:

(123) — Serginho!
— O quê?
— A bandeira está no chão.

Quando um termo da oração é assim separado por mudança de interlocutor, entende-se claramente que houve interrupção:

(124) — Serginho...
— O quê?
— ... vai fazer aniversário amanhã.

Finalmente, pode-se dizer que o vocativo tem uma resposta própria (como *O quê?* ou então *Estou aqui.* etc.), o que indica que ele pode constituir, por si só, uma frase independente. Isso não se aplica a termos individuais da oração: a oração (ou, melhor dizendo, o período) é que pode ter uma resposta, mas não o seu sujeito, ou predicado, ou adjunto circunstancial etc.

Razões como essas mostram com bastante clareza que o vocativo realmente não pertence à oração. A semântica apóia essa conclusão: o significado de um vocativo não se integra ao significado de uma oração contígua. Já com um adjunto oracional, por exemplo, há essa integração:

(125) Com franqueza, desconfio de vocês.

O significado de *com franqueza* é claramente predicado de *desconfio de vocês*. Nada de parecido se observa com o vocativo.

4 O sintagma

4.1. FUNÇÕES DE NÍVEL SUBORACIONAL

Acabamos de estudar um conjunto de funções que são desempenhadas por constituintes de nível oracional, ou seja, constituintes imediatos da oração. Esses constituintes e suas funções representam como que o arcabouço maior da oração — as grandes linhas de sua construção. Agora vamos passar ao estudo de detalhes mais finos, representados pelos elementos que formam os grandes constituintes da oração. Por exemplo, em uma oração como

(1) Esse professor é um neurótico.

temos, no nível da oração, três constituintes, a saber: o sujeito (*esse professor*); o núcleo do predicado (*é*); e o complemento do predicado (*um neurótico*). O que nos vai ocupar a partir de agora são as funções de elementos como *esse*, *professor*, *um* e *neurótico*, que são constituintes do sintagma sujeito e do sintagma complemento do predicado. Em outras palavras, vamos considerar perguntas como: Qual é a função do elemento *esse* dentro do sintagma *esse professor*?

Estudaremos em primeiro lugar um tipo especial de sintagma, denominado **sintagma nominal (SN)**. O sintagma nominal pode ser definido de maneira muito simples: é o sintagma que pode ser sujeito de alguma oração. Assim, *esse professor* é um sintagma nominal porque é sujeito da oração (1); e *um neurótico* é também um sintagma nominal porque, embora não seja sujeito em (1), pode ser sujeito em outra oração, como

(2) Um neurótico rabiscou meus livros.

Já *em São Paulo* não é um sintagma nominal, pois não poderia nunca ser sujeito de uma oração; o mesmo acontece com *imensamente rico, na semana passada, da minha tia Carolina* etc.

Vamos estudar na próxima seção a estrutura interna do sintagma nominal.

4.2. FUNÇÕES SINTÁTICAS NO SINTAGMA NOMINAL

Quando estudamos a estrutura da oração, foi possível aproveitar uma parte da análise tradicional, embora não as definições propriamente ditas; isso porque, conforme foi discutido detalhadamente em meu livro *Para uma nova gramática do português* (p. 15-20), a prática da análise segue definições implícitas (diferentes das que são dadas nas gramáticas), e essas definições implícitas são freqüentemente corretas. Assim, a noção de "sujeito" como "o termo com o qual o verbo concorda" se harmoniza com a prática da análise tradicional em quase todos os casos, embora a definição dada (e não seguida na realidade) seja "o termo do qual se afirma alguma coisa". Pode-se dizer, portanto, que a noção de sujeito proposta neste capítulo é muito próxima da noção que governa a prática tradicional de análise sintática.

No entanto, quando consideramos a estrutura interna do sintagma nominal (SN), torna-se necessário abandonar praticamente toda a análise tradicional, pois esta é excessivamente simplista e inadequada. Assim, em um SN como

(3) Aqueles seus livros de psicologia

a gramática tradicional distingue apenas duas funções: *livros* seria o "núcleo", e os demais termos (*aqueles, seus* e *de psicologia*) seriam "adjuntos adnominais". Aqui veremos que essa análise é simples demais para fazer justiça à complexidade dos fatos. Na verdade, cada um dos três "adjuntos adnominais" contidos em (3) tem um comportamento sintático diferente e, portanto, desempenha uma função diferente da de seus companheiros.

É fácil verificar isso investigando as possibilidades de posicionamento de cada um desses termos. Alguns deles têm uma posição fixa no sintagma, e não há possibilidade de permuta:

(4) a. Aqueles seus livros de psicologia desapareceram.
b. * Seus aqueles livros de psicologia desapareceram.
c. * Aqueles seus de psicologia livros desapareceram.

Verifica-se que o termo *aqueles* tem uma posição fixa nesse sintagma (só pode aparecer em primeiro lugar); e o termo *de psicologia* igualmente tem posição fixa, mas diferente (só pode aparecer em último lugar).

Por outro lado, o elemento *seus* tem possibilidade de transposição, embora limitada:

(5) a. Aqueles seus livros de psicologia desapareceram.
b. Aqueles livros seus de psicologia desapareceram.
c. ? Aqueles livros de psicologia seus desapareceram.

(Coloquei "?" no exemplo (c) porque para mim ele é de aceitabilidade marginal.)

Só com essas rapidas observações, já se pode ver claramente que os três "adjuntos adnominais" não se comportam da mesma maneira. No sintagma (3), *aqueles* e *de psicologia* têm posição fixa, mas não a mesma posição: *aqueles* precisa aparecer no início do sintagma, e *de psicologia* no final; isso já é suficiente para que recebam funções diferentes. Depois, *seus* é o único dos termos vistos que pode ser deslocado de posição. Como se vê, há complexidades dentro da estrutura do SN que são escamoteadas pela simplicidade da análise tradicional. Conseqüentemente, a análise proposta a seguir é bem mais complexa do que a tradicional. Não sinto que deva apresentar desculpas por isso: a análise de um fenômeno complexo não pode ser muito simples.

A análise que veremos é complexa no sentido em que admite um número maior de funções dentro do SN. Sabemos que a gramática tradicional só distingue, no sintagma

(3) Aqueles seus livros de psicologia

duas funções, a de "núcleo" (*livros*) e a de "adjunto adnominal" (*aqueles, seus, de psicologia*). Na nova proposta, seremos forçados a distinguir quatro, já que os três "adjuntos adnominais", como se viu, têm cada qual uma função diferente.

Quando tratamos das funções na oração, utilizamos traços de natureza diversa; assim, o traço [Ant] é de natureza posicional; já o traço [CV] exprime uma relação de concordância, que não tem nada a ver diretamente com a posição; e o traço [Q] exprime uma possibilidade de retomada pronominal. Isso foi necessário porque a estrutura da oração é relativamente livre no que diz respeito à posição dos constituintes; se tentássemos analisá-la exclusivamente em termos de posição, acabaríamos deixando de lado muitos fenômenos importantes. Por exemplo, a oposição entre adjunto circunstancial e objeto direto ficaria obliterada, pois ambos têm os mesmos privilégios de posicionamento dentro da oração.

No caso do SN, a coisa muda de figura. O SN tem uma estrutura posicionalmente muito mais rígida do que a oração; as possibilidades de mudança de ordem dos termos são poucas e bem delimitadas. Conseqüentemente, uma análise dos termos do SN em termos de posição é suficiente para revelar muitos dos grandes traços da estrutura. Aqui, portanto, vamos procurar caracterizar os termos internos do SN por meio de traços de natureza posicional — mas note-se que essa decisão não implica negar que haja (como certamente há) importantes traços da estrutura interna do SN que não têm a ver com a posição dos termos. Vamos estudar os traços posicionais de preferência, porque são importantes e porque dispomos de estudos descritivos sobre eles.

4.2.1. O SN máximo

Vamos investigar brevemente as possibilidades de ordenação do item *meu* no sintagma. Ele pode ocorrer em primeiro lugar, como em

(6) Meu computador

Mas também pode aparecer em segundo lugar, como em

(7) Aquele meu computador

Aqui já temos um problema: parece que a posição de *meu* é variável. Como exprimi-la em nossa análise?

Uma observação que nos pode ajudar é a seguinte: *meu* pode ocorrer em primeiro lugar, mas nunca quando o sintagma também tem o item *aquele*:

(8) * Meu aquele computador

Ou seja, parece que o primeiro elemento do sintagma é *aquele*; somente na sua ausência é que *meu* ocorre em primeiro lugar. Exprimimos esse fato da seguinte maneira: *aquele* ocupa sempre o primeiro lugar no SN, e *meu* o segundo; só que os "lugares" assim definidos não são obrigatoriamente preenchidos, de maneira que às vezes *meu* aparentemente ocupa a primeira posição. É como se considerássemos que existe uma primeira posição "não-preenchida" antes de *meu* em (6):

[ ] meu computador

A estrutura do SN exige a ordenação *aquele* + *meu*, mas a ocorrência de qualquer desses dois itens é opcional, de modo que qualquer dos sintagmas abaixo é aceitável:

(9) a. Computador
b. Meu computador
c. Aquele computador
d. Aquele meu computador

Já se a ordenação (que é obrigatória) for desobedecida, o resultado não será aceitável:

(8) * Meu aquele computador

Podemos, então, estabelecer como regra geral que *aquele* ocupa a primeira posição no SN, e *meu* a segunda (estou deixando de lado os itens *todos* e *ambos*, que poderiam ocorrer antes de *aqueles*; veremos em 4.3. que esses elementos ocupam uma função fora do SN). Essa regra evidentemente não vale para todos os sintagmas possíveis (já que em *meu computador* o item *meu* aparece em primeiro lugar): ela se refere a uma espécie de "SN ideal". A regra afirma, na verdade, que "quando *aquele* ocorre em um SN, ocupa a primeira posição; e quando *meu* ocorre em um SN, ocupa a segunda posição **caso a primeira esteja ocupada**". Trata-se, portanto, de uma espécie de fila de preferências para ocupar a primeira posição. É essa fila de preferências que nos servirá de base para a definição das diversas funções dentro do SN. A estrutura interna do SN se analisa a partir da pergunta: Quantas posições são possíveis, em princípio, dentro do SN?

Um sintagma em que todas as posições possíveis forem preenchidas por itens léxicos se denomina **SN máximo**. O SN máximo, na verdade, é uma abstração, porque, como veremos, não ocorre na prática; um SN máximo realizado seria tão longo e sobrecarregado que acabaria sendo rejeitado pelos falantes, por razões que nada têm a ver com a sintaxe. Seria considerado excessivamente longo e excessivamente entulhado de informação. O leitor pode julgar isso pelo sintagma

(10) Os outros dois meus mesmos velhos amigos queridos de Salvador

Esse SN, que não chega a ser máximo, é complexo a ponto de ser quase improcessável.

No entanto, o SN máximo é necessário como ponto de referência para a análise; vimos que em (6) há uma posição não preenchida (e que poderia ser preenchida por *aquele*). O mesmo raciocínio se aplica a todas as posições que compõem o SN. E a definição de cada uma das funções se fará por referência ao SN máximo.

Para tomar um exemplo, veremos adiante que o item *aquele* ocupa a sexta posição antes do núcleo do SN (estou aqui considerando apenas as posições fixas; há também posições variáveis, mas isso não precisa ser levado em conta neste exemplo); e admitamos que *amigo* é o núcleo em

(11) Aquele meu amigo

Ora, é evidente que *aquele* ocupa, em (11), a **segunda** posição antes do núcleo; o que então nos autoriza a dizer que sua posição é a sexta?

Acontece que podem ocorrer **no máximo** cinco elementos entre *aquele* e o núcleo; ou seja, pode haver cinco, quatro, três, dois, um ou nenhum:

(12) a. Aquele meu mesmo único pretenso bom amigo
b. Aquele meu mesmo único pretenso amigo
c. Aquele meu mesmo único amigo
d. Aquele meu mesmo amigo
f. Aquele amigo

(É claro que (12a), (12b) e (12c) são pouco naturais; isso se deve aos fatores mencionados acima.)

Para definir a função desempenhada por *aquele*, utilizamos sua posição de afastamento máximo, para a esquerda, do núcleo. Assim podemos distinguir *aquele* de *inesquecível*, mesmo nos sintagmas

(13) Aquele amigo
(14) Inesquecível amigo

Muito embora, nesses sintagmas, a posição de *aquele* e de *inesquecível* seja a mesma (imediatamente antes do núcleo), suas funções são diferentes, porque entre *aquele* e *amigo* poderiam ocorrer até cinco termos, mas entre *inesquecível* e *amigo* só poderia ocorrer um.

Voltemos agora aos exemplos:

(6) Meu computador

(7) Aquele meu computador

O fato de *meu* ocorrer ora na primeira, ora na segunda posição (a contar do núcleo, para a esquerda), já não precisa ser um problema. Como esses SNs não são máximos, a posição dos elementos neles não é diretamente relevante. Mas alguma coisa já pode ser observada: *aquele* tem uma função cuja posição é sempre anterior à posição de *meu*. Isso é confirmado pelo exemplo (7), e por:

(8) * Meu aquele computador

assim como por todos os sintagmas da língua que incluem os itens *meu* e *aquele*.

Em resumo, as funções no SN se definem pelas posições dos termos **em relação uns aos outros**, e não por suas posições absolutas.

4.2.2. Estrutura do SN: a área esquerda

4.2.2.1. Descrição

Para estudar a estrutura interna do SN, vamos dividi-lo em duas porções: a **área esquerda**, composta dos elementos que precedem o núcleo, e a **área direita**, composta do núcleo mais os elementos que o seguem. Vejamos primeiramente a área esquerda.

O estudo dos diversos itens que podem ocorrer na área esquerda do SN mostra uma surpreendente variedade de posicionamentos possíveis e, portanto, de funções distintas. Examinando um grande número de sintagmas, chegou-se à conclusão de que a área esquerda compreende seis **posições fixas** e quatro **posições variáveis**. As posições fixas definem seis funções, denominadas (na ordem em que ocorrem no sintagma) **determinante**, **possessivo**, **reforço**, **quantificador**, **pré-núcleo externo** e **pré-núcleo interno**. Abreviando, temos, respectivamente, **Det**, **Poss**, **Ref**, **Qf**, **PNE** e **PNI**. As posições variáveis ocorrem nos intervalos entre as posições fixas, exceto entre os dois pré-núcleos, onde não pode ocorrer nenhum item. O esquema completo da área esquerda é o seguinte (as posições variáveis são marcadas como "PV"):

[Det PV4 Poss PV3 Ref PV2 Qf PV1 PNE PNI]

Cada posição define, portanto, uma função; diz-se de *aquele* em (7) que é um determinante, porque ocupa a primeira posição. No entanto, as quatro posições variáveis definem uma única função, a de **numerador** (Num), que tem a propriedade de ocorrer nas quatro posições marcadas "PV" no esquema. Isso se deve ao fato de que os mesmos itens léxicos ocorrem nessas quatro posições, sendo as versões resultantes totalmente

correspondentes. Temos, portanto, dez posições, mas apenas sete funções na área esquerda.

Essas dez posições precedem o núcleo no SN máximo. Elas se definem de maneira simples: o Det ocupa a primeira posição no SN máximo, o Poss a segunda, o Ref a terceira e assim por diante.

No entanto, a dificuldade de se obter SNs máximos faz com que essas definições sejam relativamente pouco úteis. Assim, a depreensão das diversas funções teve de ser feita por meio do exame de grande número de SNs menores. Por exemplo, observou-se que nenhum elemento do SN pode ocorrer antes de *aquele* (relembro que *todos, ambos* são externos ao SN; ver 4.3.). Assim, *aquele* deve ocupar a primeira posição em qualquer SN — e, portanto, também a primeira no SN máximo.

Já a palavra *meu* e os possessivos em geral podem ocorrer logo após *aquele,* o que sugere que *meu* ocupa a segunda posição. Na verdade, há dois itens que podem ocorrer entre *aquele* e *meu,* e que poderiam ser considerados como ocupando a segunda posição (*meu* passaria para a terceira): são os itens *outro* e *dois* (e os cardinais em geral):

(15) a. Aquele meu sapato
b. Aquele outro meu sapato
c. Aqueles dois meus sapatos

Apesar disso, analiso *meu* como ocupando a segunda posição (isto é, a função de Poss), porque *outro* e *dois* ocorrem em várias outras posições no SN. Podem, por exemplo, aparecer após *meu:*

(16) a. Aquele meu outro sapato
b. Aqueles meus dois sapatos

São apenas esses dois itens que podem ocorrer dessa forma em várias posições relativamente aos demais componentes do SN. Por isso, criou-se para eles uma função especial, a de **numerador**, que se define pela possibilidade de ocorrer livremente em quatro posições no SN: PV4, PV3, PV2 e PV1. Conclui-se, então, que *meu* pode ter a função de possessivo; *aquele* de determinante; e *dois* e *outro* de numerador.

Não seria possível reproduzir aqui todo o raciocínio que levou ao estabelecimento da seqüência completa de funções da área esquerda do SN; o leitor interessado poderá reportar-se ao relatório do projeto TENPo (Perini et al., em preparação). Aqui limitar-me-ei a mostrar os resultados, dar uns poucos exemplos e comentar alguns problemas.

Os itens que desempenham as funções de Det, Poss, Ref, Qf, PNI e Num são pouco numerosos, pertencendo a classes fechadas; dou abaixo uma lista razoavelmente completa desses itens. Já aqueles que podem ter a função de PNE são muito numerosos e constituem uma classe aberta; a lista dada aqui é, portanto, apenas uma pequena amostra:

Função	Itens que podem desempenhá-la
Det	*o, este, esse, aquele, algum, nenhum, um.*
Poss	*meu, seu, nosso* etc.
Ref	*mesmo, próprio, certo.*
Qf	*poucos, vários, diversos, muitos, único, primeiro (segundo, terceiro* etc.*).*
PNE	*mero, pretenso, meio, suposto, reles, inesquecível, ilusório, simples, bom, velho, novo* etc. [classe aberta]
PNI	*mau, novo, velho, claro, grande.*
Num	*outro, dois (três, quatro* etc.*).*

NOTAS

(a) O item *meu* representa todos os possessivos; *primeiro* todos os ordinais; *dois* todos os cardinais, exceto *um*, que tem comportamento diferente.

(b) Note-se que os itens que podem ser PNI também podem ser PNE, o que naturalmente dificulta a análise. Por outro lado, há itens que só podem ser PNE: *mero, pretenso, meio, suposto* e talvez *reles* (*reles* só muito raramente ocorre como modificador, isto é, após o núcleo).

(c) Alguns itens foram citados no plural, por serem relativamente raros no singular.

(d) Faltam na lista alguns itens cuja função é obscura; ver 4.2.2.2.

Uma observação final: de todas essas funções, a única que pode ocorrer mais de uma vez no mesmo sintagma é a de Num. Naturalmente, o Num só pode ocorrer duas vezes, já que apenas dois itens o podem ocupar; e, quando ocorrem os dois, a ordem de um relativamente ao outro é livre:

(17) a. Os meus outros dois sapatos
b. Os meus dois outros sapatos

4.2.2.2. Itens de função duvidosa

Alguns dos itens da área esquerda do SN são de função duvidosa, porque é difícil obter dados relevantes para sua análise. Isso se deve principalmente às restrições de compatibilidade entre os diversos itens; vejamos, então, primeiro o que vêm a ser essas restrições.

Por razões tanto sintáticas quanto semânticas, certos itens da área esquerda não podem coocorrer no mesmo sintagma com certos outros itens. Assim, não podemos jamais colocar lado a lado itens que desempenham a mesma função (a menos que esta seja a de numerador):

(18) * O aquele sapato

Isso deriva automaticamente do princípio de que tais funções não são repetíveis.

Mas há outras restrições, de conseqüências mais sérias. Assim, nunca podemos colocar no mesmo sintagma os itens *dois* e *poucos*:

(19) * As duas poucas garotas

Isso se deve, evidentemente, a uma incompatibilidade semântica entre os itens: *poucos* denota uma quantidade pequena, mas indeterminada; *dois* denota uma quantidade exata. Logo, usar ambos para qualificar a mesma entidade é contraditório.

Em outros casos, a razão da incompatibilidade é menos evidente; acredita-se que as incompatibilidades são semânticas na maioria das vezes, mas é provável que haja também incompatibilidades de ordem formal. O problema não mereceu ainda estudos que o esclareçam.

No que nos interessa, porém, as conseqüências são claras: quanto mais restrições de compatibilidade um item sofre, mais difícil é determinar sua função no SN, pois esta depende de sua ordenação com relação aos demais itens. Assim, chegamos a casos extremos, como o dos itens *cada* e *todo* (não o item *todos*, que varia em gênero e número; aqui tratamos do item *todo*, que só varia em gênero e pode ocorrer sem determinante: *todo homem*). Esses itens são incompatíveis com a maioria dos outros itens da área esquerda; coocorrem apenas com o PNE e com o PNI e, nesse caso, vêm antes deles:

(20) Todo bom professor

(21) Cada inesquecível viagem

Cada (mas não *todo*) coocorre também com os cardinais, como em

(22) Cada três recibos dão direito a um carro zero km.

Mas como sabemos que os cardinais são de posição variável, isso não nos diz muito sobre *cada*. A única coisa que se pode concluir, então, é que *todo* e *cada* não podem ser PNE nem PNI; sua análise completa deverá, no mais, ficar indeterminada por indisponibilidade de dados.

Outro item que apresenta o mesmo problema de baixa compatibilidade, e que portanto é de análise difícil, é *qualquer*; deste também se sabe que não pode ser PNE nem PNI, e pouco mais.

Também surgem dúvidas de análise com os itens, já mencionados, que podem desempenhar mais de uma função. Na área esquerda, estes são os itens que podem ser PNI (*bom, mau, novo, velho, claro, grande*), que também podem ser PNE. Como esses mesmos itens (e alguns outros poucos da área esquerda) podem igualmente ocorrer na área direita, sua discussão ficará para a seção seguinte.

4.2.3. A área direita

Passemos agora à área direita do SN, que compreende o núcleo e os termos que ocorrem depois dele. A descrição oferecida aqui é bem

menos segura do que a vista acima para a área esquerda, porque a pesquisa detalhada está apenas começando. A área dos modificadores é especialmente mal compreendida no momento, e é possível que seja necessário distinguir mais de dois modificadores. Isso deverá ser esclarecido por pesquisas, já em andamento. No entanto, pelo menos os grandes traços já se deixam ver, e é possível dar uma visão geral dos termos que compõem a área direita.

4.2.3.1. Descrição

Distinguem-se na área direita três funções: **núcleo do SN** (NSN), **modificador interno** (ModI) e **modificador externo** (ModE). Esses termos ocorrem nessa ordem e podem ser definidos, respectivamente, como o antepenúltimo, o penúltimo e o último termo no SN máximo (ou, alternativamente, como o 7º, 8º e 9º termo, não se contando as posições variáveis da área esquerda). Um exemplo é

(23) Um ataque cardíaco fulminante

Um, como já vimos, é um Det. *Ataque* é o NSN, *cardíaco* o ModI e *fulminante* o ModE. A necessidade de propor dois modificadores vem do fato de que sua ordenação é fixa:

(24) * Um ataque fulminante cardíaco

A estrutura, como se vê, é mais simples do que a da área esquerda. No entanto, a análise, na prática, é mais difícil, porque encontra dois obstáculos sérios: o alto grau de polivalência dos itens envolvidos e o fato de que estes pertencem a classes abertas.

Um dos fatores que facilitam a análise na área esquerda é que a maioria das funções são ocupadas por itens especializados; assim, *nosso* só pode ser Poss, *aquele* só pode ser Det etc. Isso nos fornece uma série de pontos de referência para analisar sintagmas. Essa facilidade, entretanto, é muito diminuída na área direita: embora haja também aí itens especializados, a norma é a polivalência funcional. Conseqüentemente, as dúvidas de análise são muito mais freqüentes.

Outro fator de facilitação de que desfruta a área esquerda é o pequeno tamanho das classes envolvidas. Assim, a função de Poss pode ser ocupada apenas por cinco itens: *meu*, *teu, seu, nosso, vosso.* Isso se repete para todas as posições, exceto a de PNE. Novamente, a situação é diferente na área esquerda: todas as três funções são ocupadas por classes abertas, inviabilizando a confecção de listas de itens; não é possível (pelo menos em uma gramática) dar uma lista completa dos itens léxicos que podem ser núcleos de um SN, ou modificadores internos ou externos; e, além do mais, as funções de NSN e de ModE podem ser preenchidas por sintagmas maiores, e não apenas por palavras individuais.

Esses fatos nos deixam sem critérios cômodos de análise para a área direita; seremos então forçados a examinar cada caso particular, lançando mão de pistas mais ou menos decisivas; é o que faremos nos exemplos a serem estudados na seção 4.2.3.4.

4.2.3.2. Justificação das funções

Antes disso, vou expor as razões pelas quais é necessário distinguir pelo menos três funções na área direita: NSN, ModI e ModE.

Tomemos o sintagma

(23) Um ataque cardíaco fulminante

Já sabemos que os três últimos elementos estão estritamente ordenados, não havendo em geral liberdade de transposição:

(25) a. * Um ataque fulminante cardíaco
b. * Um cardíaco ataque fulminante
c. * Um cardíaco fulminante ataque
d. * Um fulminante cardíaco ataque

A única transposição aceitável nos dará

(26) Um fulminante ataque cardíaco

Esses dados autorizam algumas conclusões, a saber:

(a) *cardíaco* tem função diferente de *fulminante*, pois *cardíaco* não pode ocorrer nem em último lugar, nem antes de *ataque*, ao passo que *fulminante* pode ocorrer em ambas essas posições;

(b) *ataque* e *cardíaco* têm igualmente funções diferentes, pois só podem ocorrer na ordem *ataque cardíaco*, e não o inverso; se tivessem a mesma função, ambas as ordens deveriam ser possíveis;

(c) *ataque* e *fulminante* também têm funções diferentes, pois *ataque*, mas não *fulminante*, pode ocorrer logo antes de *cardíaco*.

Ou seja, os três itens, *ataque*, *cardíaco* e *fulminante*, desempenham três funções distintas no SN (23).

A análise de (23), naturalmente, ainda não está pronta: sabemos que cada um dos itens (exceto *um*, que já sabemos pertencer à área esquerda, com a função de Det) tem uma função. Mas que funções serão essas?

Vamos admitir que o núcleo do SN tem a propriedade de constituir, por si só, um SN — isso estaria em consonância com a noção habitual de "núcleo" como um elemento essencial. Podemos incluir essa exigência como parte da definição de NSN. Ora, dos três itens em exame, apenas *ataque* poderia ocorrer sozinho como um SN:

(27) Ataque só se trata em hospital.

Note-se que *fulminante* ou *cardíaco* nunca poderiam ocorrer assim. A restrição se estende à coocorrência com os itens da área esquerda: so-

mente *ataque* pode ocorrer nessa situação:

(28) Um ataque / Esse ataque / Seu ataque
(29) * Um fulminante / * Esse fulminante / * Seu fulminante
(30) * Um cardíaco / * Esse cardíaco / * Seu cardíaco

Podemos atribuir, portanto, a *ataque* a função de núcleo do SN (23). Isso nos fornece a análise de *cardíaco* (ModI) e *fulminante* (ModE). A diferença de funções entre *cardíaco* e *fulminante* é confirmada pelo seguinte fato: só *fulminante*, e não *cardíaco*, é que pode ser separado do resto do SN por algum sinal de pontuação:

(31) a. Um ataque cardíaco, fulminante
b. Um ataque, fulminante
c. * Um ataque, cardíaco, fulminante
d. * Um ataque, cardíaco

Ou seja, podemos acrescentar à definição de modificador externo a propriedade de ser separável do resto do SN por sinal de pontuação.

Resta considerar *fulminante* em

(26) Um fulminante ataque cardíaco

Vamos analisá-lo como PNE. Essa função, como vimos, é desempenhada por uma classe aberta; e boa parte dos itens que podem ocorrer como modificadores podem também ocorrer como PNE. É isso que explica a ocorrência de pares de SNs como

(32) a. Uma experiência fantástica
b. Uma fantástica experiência
(33) a. Um almoço rápido
b. Um rápido almoço

Os exemplos (a) não são sinônimos perfeitos dos exemplos (b); isso se pode atribuir justamente à diferença de função de *fantástica* e *rápido*, que é levada em conta pelas regras que interpretam esses sintagmas.

O sintagma (23), portanto, analisa-se como uma seqüência de Det + NSN + ModI + ModE. E duas informações importantes, que nos serão úteis depois, são: (a) o NSN é o único desses termos que pode constituir um SN seja sozinho, seja precedido de termo da área esquerda; os modificadores (como, aliás, também o PNE e PNI) só podem ocorrer no SN na presença de um núcleo; e (b) dos modificadores, somente o externo pode ser separado do resto do SN por sinal de pontuação.

4.2.3.3. Discussão

Dada a carência de um critério decisivo para estabelecer, na prática, a função dos constituintes da área direita, torna-se importante acumular pistas que nos possam orientar na análise de casos particulares. A mais importante dessas pistas é a identificação de itens especializados, isto é,

aqueles que só podem ter uma função dentro do sintagma.

Como exemplo, vamos examinar a palavra *ataque*. Analisando (23), vimos que *ataque* é NSN; isso foi apurado verificando-se que *cardíaco* e *fulminante* não podem ser NSN. Seguindo o mesmo raciocínio, pode-se identificar os itens que não podem ser NSNs. Agora vou mostrar que *ataque*, quando dentro de um SN, só pode ser núcleo: trata-se de um item especializado.

Começarei mostrando que *japonês* pode ser núcleo de um SN: pode ocorrer no SN sozinho ou, então, acompanhado de um termo da área esquerda:

(34) Japonês geralmente é moreno.

(35) Minha neta está namorando um japonês.

Como sabemos, isso mostra que *japonês* é núcleo dos SNs *japonês* e *um japonês*.

Agora tomemos a frase

(36) O ataque japonês se deu de madrugada.

O sintagma *o ataque japonês* (deixando de lado *o*) pode ter, em princípio, uma das seguintes análises:

(a) NSN ModI
(b) PNE NSN
(c) PNE ModI
(d) ModI ModE
(e) NSN ModE

As análises (d) e (e) podem ser afastadas porque *japonês* não pode ser separado por vírgula e, portanto, não deve ser ModE:

(37) * O ataque, japonês, se deu de madrugada.

Note-se a diferença se *japonês* é substituído por um item que pode ser ModE:

(38) O ataque, inesperado, se deu de madrugada.

As alternativas (b) e (c) dependem de se analisar *ataque* como PNE. Mas nesse caso é de se esperar que possa ocorrer um PNI entre *ataque* e *japonês*, desde que não haja choque semântico. Mas isso não é possível:

(39) * O ataque novo japonês

Em vez disso, o presumível PNI ocorre antes de *ataque*:

(40) O novo ataque japonês

O sintagma (40) mostra que não há choque semântico entre *novo* e *ataque japonês*; portanto, a má formação de (39) deve ser de origem sintática. Essa observação é compatível com a hipótese de que *ataque* não é PNE; somos forçados a rejeitar as análises (b) e (c). Resta-nos a análise (a), segundo a qual *ataque* é NSN, e *japonês* é ModI. Podemos tomá-la como correta.

Até o momento, mostramos apenas que *ataque* é NSN em (36). Mas sabemos que *japonês* pode igual-

mente ser NSN; se *ataque* pudesse ser ModE (ou ModI), deveria ser possível montar um sintagma com *japonês* como NSN (pois esse item pode ser NSN) e *ataque* como modificador. No entanto, isso não é possível:

(41) * O japonês ataque

A conclusão é que ***ataque* só pode ser NSN**. Seguindo esse tipo de raciocínio, poderemos elaborar uma lista de itens especializados na função de NSN, o que nos dará um ponto de apoio para a análise de muitos sintagmas.

Pode-se também mostrar que o item *cardíaco* só pode ser ModI. Em primeiro lugar, *cardíaco*, como já vimos, não pode ser NSN. Também não pode ser ModE, porque nunca se separa por sinal de pontuação de seu SN. Finalmente, não pode ser nem PNE nem PNI, porque não pode ocorrer no SN antes do núcleo; ver (25) e

(42) * Um cardíaco ataque

Uma lista dos itens que só podem ser ModI também nos será útil como auxiliar na análise.

Desse modo, podemos elaborar listas de itens especializados para cada uma das funções do SN. A única função que certamente não tem itens especializados é o PNI: todas as palavras conhecidas que podem ser PNI podem também ser modificadores. Quanto ao ModE, ainda não se fez uma investigação, mas é provável que haja formas especializadas (palavras e/ou sintagmas preposicionados, como *do Brasil*, *de confiança* etc.).

4.2.3.4. Exemplificação

Vejamos agora a análise de alguns sintagmas que podem apresentar dificuldade. Esses mesmos sintagmas foram analisados em um livro anterior (Perini, 1989, p. 171-8); mas, como a análise do SN foi modificada, com o acréscimo de diversas funções, vale a pena discuti-los de novo.

Vejamos primeiramente o sintagma

(43) Um velho

A palavra *velho*, ao que tudo indica, pode desempenhar pelo menos três funções no SN: pré-núcleo interno, núcleo e modificador externo. Talvez possa ser ainda modificador interno ou pré-núcleo externo, mas vamos deixar de lado essas possibilidades, porque ainda não foram devidamente investigadas.

Como já sabemos que nem o PNI nem o ModE podem ocorrer no SN acompanhados apenas de um elemento da área esquerda, podemos concluir desde já que *velho* em (43) é NSN.

Já em

(44) Um velho palhaço

a função de *velho* não é imediatamente evidente.

Primeiro, é preciso observar que (44) tem duas interpretações semânticas (ou seja, é ambíguo): um "palhaço que é velho", ou um "velho que é palhaço". Como veremos, essa ambigüidade decorre de uma dualidade de estruturas. Em outras palavras, (44) é a representação explícita única de dois sintagmas nominais distintos.

A palavra *palhaço* não pode ser PNI nem PNE. Mostra-se isso tentando colocá-la antes de qualquer palavra que só possa ser NSN, como, por exemplo, *crocodilo*:

(45) * Um palhaço crocodilo

ao passo que *palhaço* pode ocorrer como modificador (após o núcleo) em

(46) Um crocodilo palhaço

Portanto, *palhaço* em (44) pode ser NSN ou então ModE. E, na verdade, é as duas coisas, porque representa dois SNs homônimos, mas distintos. Um deles se analisa como Det + PNI + NSN e significa "um palhaço idoso"; o outro se analisa como Det + NSN + ModE e significa "um velho que é ou se comporta como palhaço".

Passemos agora a

(47) O imperador menino

Todo SN tem núcleo, como se verá em 4.4.; portanto, as análises possíveis desse sintagma são as seguintes:

(a) NSN NSN
(b) PNI NSN
(c) PNE NSN
(d) NSN ModE
(e) NSN ModI

A análise (a) deve ser afastada, porque a evidência disponível mostra que o NSN não pode ocorrer repetido no mesmo sintagma. Sabemos isso porque nunca se pode colocar juntas no SN duas ou mais das palavras que só podem ser SNs, como *ataque*. A análise (b) também deve ser excluída, porque a classe dos itens que podem ser PNI é muito limitada, e certamente *imperador* não é um deles.

Quanto a (c), não deve ser correta, porque é possível antepor um PNE a *imperador menino*:

(48) Um mero imperador menino

Também (d) não pode ser correta, porque não se pode separar *menino* por vírgula.

Portanto, a análise de (47) é a definida por (e): Det + NSN + ModI.

4.2.4. Repetições de termos no SN

Ao que parece, os únicos termos do SN que podem ocorrer repetidos são o Num e os modificadores ou, pelo menos, o modificador ex-

terno. Vimos exemplos de repetição de Num na seção 4.2.2.1.: *os meus outros dois sapatos*, onde tanto *outros* quanto *dois* têm a função de Num. Quanto aos modificadores, é difícil definir com exatidão suas possibilidades de repetição, porque a área é ainda mal conhecida. Mas sintagmas como

(49) O livro de exercícios verde do Rogério

nos podem dizer alguma coisa. Como *livro* não pode ser PNI nem PNE, deve ser o núcleo. Por conseguinte, temos três modificadores, a saber, *de exercícios*; *verde*; *do Rogério*. Se só houver (como parece) duas funções após o NSN (ModI e ModE), pelo menos uma delas estará repetida em (49).

Isso é confirmado pela possibilidade de permuta de ordem em (49):

(50) O livro verde de exercícios do Rogério

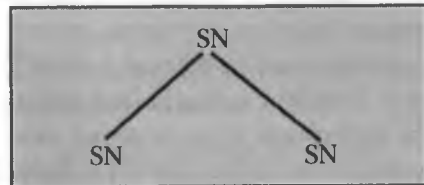
A permuta sofre restrições que ainda não são bem compreendidas:

(51) * O livro do Rogério de exercícios verde

Como a área dos modificadores é ainda um tanto misteriosa, vou deixar a questão por aqui, à espera que a pesquisa a venha esclarecer.

Antes de encerrar, porém, devo chamar a atenção do leitor para o seguinte: a **coordenação** não vale como repetição. Muitos termos do SN podem ser coordenados, como, por exemplo, o NSN:

(52) Antônio e Maria

Isso não significa que (52) seja um SN com dois núcleos. O que temos ali é a coordenação de dois SNs, formando por sua vez um terceiro SN, o que pode ser indicado assim:

As coordenações serão estudadas na seção 5.2.2.

4.2.5. Sobre as incertezas da análise

Creio que algumas pessoas se sentirão frustradas com a situação apresentada nas páginas precedentes: tivemos que conformar-nos com o fato de que a análise de certos termos só pode ser verificada por meio de uma abordagem indireta. E há pior: em certos casos, a análise fica simplesmente em suspenso, por falta de pistas relevantes. Como se compara essa situação com a da gramática tradicional, que parece prometer, implicitamente, uma decisão segura para cada caso?

A diferença está no que se entende por segurança. A prática tradicional obriga o estudioso a aceitar constantemente análises que ele próprio não saberia justificar e que freqüentemente vão de encontro à teoria explicitada: é como se o objetivo da atividade gramatical fosse obter uma resposta para tudo, em vez de procurar a compreensão do fenômeno estudado. O resultado é que se escamoteiam as incertezas, mascarando-as como problemas já resolvidos. Trabalhar com a análise sintática tradicional dá geralmente uma sensação de "coisa pronta"; é muito raro encontrar-se um ponto que precise ser mais pesquisado, precisamente porque tais pontos são escondidos. Não é de admirar que a disciplina não atraia a curiosidade dos novos pesquisadores.

Mas nenhum aspecto do mundo, e menos que todos as línguas naturais, é cabalmente conhecido. O valor de um trabalho científico se avalia não apenas em termos dos problemas que resolve, mas também em termos das questões que levanta. Uma boa análise resolve alguns pontos e ajuda a formular as perguntas cruciais que orientarão pesquisas futuras. Ou seja, as incertezas são um fato do dia-a-dia do gramático e devem ser encaradas com espírito positivo: não como um motivo de desânimo, mas como um convite à pesquisa.

4.3. O PREDETERMINANTE

Vamos agora examinar a função do elemento *todos* no sintagma *todos os crocodilos*:

(53) Todos os crocodilos gostam de frango.

Se *todos os crocodilos* for um SN, teremos aqui um elemento, *todos*, que ocorre **antes** do determinante — contrariando, assim, a análise dada acima do determinante como primeiro elemento do SN. No entanto, há razões para suspeitar que *todos* não pertence ao SN: ao contrário dos membros integrantes do SN, ele pode ser transportado para posições não-contíguas ao que seria o restante de seu SN, como em

(54) Os crocodilos gostam todos de frango.

(Essas duas frases se correspondem, razão pela qual se pode falar de "transporte" de *todos*.)

Por isso, tem sido proposto que *todos* seja considerado um elemento externo ao SN, que seria, portanto, reduzido, em (53), a *os crocodilos*; vou adotar essa análise aqui. A função de *todos* se denominará **predeterminante** (PDet), um termo que não pertence à estrutura do SN. (Neste particular, estou modificando a proposta contida em meu livro de 1989, onde analisei o predeterminante como o primeiro elemento do SN.)

A análise do predeterminante apresenta, como se vê, o problema especial de determinar as posições em que ele pode ocorrer na oração (preservando a relação de correspondência e também, evidentemente, a relação semântica com o SN). Veremos a seguir que o PDet tem uma liberdade de movimentação bastante grande dentro da oração e que as restrições a essa movimentação são em geral de caráter semântico.

O PDet pode ocorrer imediatamente antes do SN a que se relaciona, como em

(53) Todos os crocodilos gostam de frango.

Pode ocorrer também logo após o SN:

(55) Os crocodilos todos gostam de frango.

Uma observação interessante é que muitos falantes, embora aceitem (55) como bem formada, "sentem" que *os crocodilos todos* não forma um constituinte — o que pode ser tomado como uma indicação a mais de que *todos* não pertence ao SN.

Além disso, como vimos, o PDet pode aparecer também separado do SN, embora a correspondência e a relação semântica se mantenham. Assim, temos

(54) Os crocodilos gostam todos de frango.

Aqui, o PDet se coloca logo após o NdP. Mas ele pode ocorrer também logo após o auxiliar:

(56) Os motoristas estão todos votando pela greve.

Mesmo quando há auxiliar, o PDet ainda pode ficar depois do NdP:

(57) Os motoristas estão votando todos pela greve.

Desse modo, podemos resumir os fatos observados até o momento da seguinte maneira:

Posições do PDet na oração
(1ª versão)

— *logo antes do "seu" SN;*
— *logo depois do "seu" SN;*
— *logo após o Aux;*
— *logo após o NdP.*

Essa formulação tem diversos defeitos. Primeiro, é necessário especificar o "seu" SN — isto é, o SN com o qual o PDet se relaciona. Essa relação, como sabemos, é de natureza semântica (embora possivelmente também inclua um componente sintático). Depois, ficam sem explicação alguns casos como o de

(58) * Maria está todos vendendo os móveis.

Pela formulação dada, deveria ser possível relacionar *todos* com *os móveis*, mas isso não é possível; conseqüentemente, a frase é mal formada, já que não se pode relacionar *todos*

com *Maria*, por razões semânticas óbvias.

Tendo a acreditar que as restrições sintáticas ao posicionamento do PDet na oração são muito simples; a maior parte do trabalho deverá ser feito por condições semânticas.

A sintaxe só precisa, na verdade, estipular o seguinte:

Posições do PDet na oração
(2ª versão)

O PDet pode colocar-se na oração entre quaisquer dois termos de nível oracional.

Assim, a única restrição sintática é a impossibilidade de se inserir o PDet dentro de algum termo oracional, como em

(59) * Os crocodilos fugiram do todos zoológico.

onde *todos* se insere, indevidamente, dentro do constituinte oracional *do zoológico.*

A segunda versão ainda precisa ser completada devido a um caso particular em que o PDet se coloca **dentro** de um SN. Um exemplo é

(60) Os motoristas todos do Brasil entraram em greve.

Tudo indica que *os motoristas todos do Brasil* é um SN; no entanto, *todos,* aparentemente um PDet, está colocado no interior desse SN.

O PDet, em tais casos, se coloca antes do modificador. No entanto, nem todo modificador aceita essa inserção, como se vê em

(61) * Os motoristas todos honestos entraram em greve.

É como se tivéssemos dois tipos de modificadores, dos quais apenas um admite a anteposição do PDet. Essa impressão é reforçada pelo fato de que, quando ocorrem ambos os tipos de modificadores no mesmo sintagma, o modificador que admite anteposição de PDet ocorre obrigatoriamente em segundo lugar:

(62) Os motoristas honestos do Brasil...

(63) * Os motoristas do Brasil honestos...

Trata-se, aparentemente, de uma diferença de comportamento entre o modificador interno e o externo. Podemos dizer que o PDet pode colocar-se logo antes do ModE, mas não antes do ModI — o que, aliás, coaduna-se com o que sabemos da maior facilidade de separação que existe entre o ModE e o NSN.

Isso estabelecido, podemos formular a versão final da estipulação sintática quanto à posição do PDet:

Posições do PDet na oração
(versão final)

O PDet pode colocar-se na oração:

(a) dentro de um SN, logo antes de um ModE; ou

(b) entre quaisquer dois termos de nível oracional.

Isso é tudo o que a sintaxe tem a dizer quanto ao posicionamento do PDet.

Um grande número de fatos fica, porém, sem explicação. Por exemplo, por que não se pode dizer

(58) * Maria está todos vendendo os móveis.

já que essa frase não desobedece à restrição dada acima? A resposta é que esse exemplo, e vários outros, deverá ser excluído por restrições semânticas; em outras palavras, o que está errado com (58) não é a estrutura sintática, mas a interpretação semântica. As restrições semânticas são descritas no Quadro 10-B, ao qual remeto o leitor.

4.4. AINDA O SN: SNs SEM NÚCLEO?

Antes de encerrar o estudo do sintagma nominal, vamos considerar o problema da análise de certos casos em que se poderia defender que o SN ocorre sem NSN. Por exemplo, sejam as frases

(64) Alguns pensam que educação é escola.

(65) Muitos gostam de uísque com guaraná.

(66) Todos desconfiam de Sueli.

O problema é se vamos analisar, digamos, o termo *muitos* em (65) como quantificador, que é a sua função habitual, ou como NSN, já que ele constitui, por si só, um SN. Vou defender a idéia de que se trata de um NSN, o que acarreta duas generalizações de certo interesse: (a) todo SN tem núcleo; (b) certos itens da área esquerda podem também ocorrer como NSN — ou seja, não há uma relação biunívoca total entre classes e funções na área esquerda.

Isso quer dizer, em outras palavras, que, encontrando-se um item como *muitos*, não se pode analisá-lo automaticamente como quantificador. Por outro lado, mantém-se a generalização de que, conhecendo a função na área esquerda (exceto o PNE), pode-se fazer a lista dos itens que a desempenham. E também continuam a existir itens totalmente especializados na área esquerda, embora já não sejam todos: por exemplo, *o, mesmo, próprio* não podem ser NSNs.

Vejamos o caso de *alguns* em:

(64) Alguns pensam que educação é escola.

Poderíamos analisar *alguns* como determinante, deixando assim o SN sem núcleo. Vou tentar mostrar que essa análise traz certos inconvenientes.

Observemos primeiro que a propriedade de ocorrer sozinho no SN não vale para todos os elementos que podem ser determinantes. Por exemplo, *o, este, aquele* só podem constituir SNs em situações

anafóricas (ver 2.4.). É preciso, portanto, marcar essa diferença entre *alguns* e os outros itens. Se *alguns* for Det em (64), teremos de criar um novo traço para diferençá-lo, especificando que pode ocorrer sozinho no SN. Trata-se de um traço peculiar, pois não marca propriamente a função que o item pode desempenhar, mas a companhia que ele pode ter no sintagma. Note-se que tanto *alguns* quanto os outros itens serão marcados, segundo essa hipótese, como podendo ser Det e NSN.

Já, se analisarmos *alguns* como NSN em (64), a diferença entre *alguns* e os demais itens será expressa da seguinte maneira: *alguns* pode ser Det e NSN, e *o*, *este* etc. podem ser Det, mas não NSN. Não será necessário criar um traço especial para exprimir a idiossincrasia de *alguns*. O mesmo argumento vale, obviamente, para *muitos* e *todos*. Temos aqui, portanto, um argumento em favor de admitir a presença de NSN nas frases (64)-(66).

Um segundo argumento deriva do exemplo:

(66) Todos desconfiam de Sueli.

Existe aqui, evidentemente, uma relação de concordância entre *todos* e *desconfiam*. Mas, se *todos* for predeterminante, então não será parte de nenhum SN — conseqüentemente não haverá SN sujeito em (66). Isso complica nossa concepção de concordância verbal: a concordância pode fazer-se com um elemento que não é um SN, mas só em certos casos especiais, a saber, quando há um PDet e não há candidato a sujeito.

Por outro lado, se *todos* for NSN, teremos um SN (e, aliás, nenhum PDet), e a concordância funcionará da maneira habitual. Temos aqui um segundo argumento em favor de postular um NSN nas frases em pauta.

Portanto, analisaremos essas frases como contendo cada uma um SN sujeito; em (64), (65) e (66), esse sujeito é, respectivamente, *alguns*, *muitos* e *todos*. O item *alguns* pode desempenhar as funções de Det e NSN; *muitos*, de Qf e NSN; e *todos*, de PDet e NSN. Já *o*, por exemplo, pode ser Det, mas não NSN, e assim por diante. Por razões de simetria e simplicidade, é claro que o mesmo raciocínio se aplica igualmente quando não se trata de sujeito, como em

(67) Minha decisão desagradou a muitos.

Adotaremos a mesma análise para a construção (um tanto rara) exemplificada em

(68) Fiz um seguro para proteger os meus.

Vamos admitir que *meus* é NSN em *os meus*. Aqui a situação não é idêntica à vista em (67), e as coisas não são tão claras; no momento, porém, essa me parece a melhor solução.

4.5. FUNÇÕES SINTÁTICAS NO SINTAGMA ADJETIVO

4.5.1. O SAdj

Há pelo menos duas outras classes importantes de sintagmas, além do SN: o sintagma adjetivo (SAdj) e o sintagma adverbial SAdv). Nesta seção, vamos estudar o SAdj; teremos de contentar-nos com uma descrição menos detalhada do que a do SN vista na seção anterior, porque o SAdj não está tão bem estudado. De qualquer modo, já é possível dar uma visão geral de sua estrutura.

Podemos definir o SAdj como a classe dos constituintes que podem desempenhar a função de modificador interno ou modificador externo. Assim, são sintagmas adjetivos as porções em itálico dos exemplos abaixo (o SN, em cada caso, está entre colchetes):

(69) Vovô é [um homem *honesto*].
(70) Vovô é [um homem *muito honesto*].
(71) Vovô é [um homem *muito satisfeito com a vida*].

Cada um dos SNs entre colchetes compreende um determinante (*um*), um núcleo (*homem*) e um modificador (a parte em itálico). O modificador é sempre preenchido por um sintagma adjetivo.

O fato de que o modificador seja sempre preenchido por um SAdj nos fornece uma definição conveniente de SAdj. Mas o SAdj pode também desempenhar outras funções: pode ser complemento do predicado, como em

(72) Vovô está *muito satisfeito com a vida.*

ou então predicativo, como em

(73) Nós o consideramos *muito honesto.*

Essas são as três funções que um SAdj pode ocupar. Rigorosamente falando, pois, SAdj é um constituinte que pode ser modificador, complemento do predicado ou predicativo.

4.5.1.1. Estrutura interna do SAdj

A estrutura interna do SAdj encerra alguns mistérios, que estão ainda à espera de estudos aprofundados. Vou aqui apenas apontar algumas linhas gerais, pois qualquer tentativa de análise mais completa seria prematura. Mesmo o estabelecimento de um SAdj máximo apresenta, como veremos, certas dificuldades.

Uma coisa que se pode dizer é que o SAdj compreende, entre outros, dois termos exemplificados em

(74) Vovô é um homem *satisfeito com a vida.*

O SAdj (em itálico no exemplo) consta de dois termos, *satisfeito* e *com a vida*, que não podem trocar de posição sem afetar a unidade do sintagma:

(75) * Vovô é um homem com a vida satisfeito

(76) ? Vovô é um homem, com a vida, satisfeito.

Este último exemplo pode ser aceito se se subentender um forte componente contrastivo, como em

(77) ... com a vida, satisfeito; mas com a morte, inconformado.

De qualquer modo, aí já não se pode falar de um sintagma unificado, de modo que vamos deixar tais casos fora da discussão.

Dos dois termos exemplificados em (74), o segundo (*com a vida*) tem as seguintes características: (a) aparece sempre em último lugar no sintagma, e (b) é sempre composto de preposição seguida de SN ou, então, de uma das palavras formadas com o sufixo *-mente* e que tradicionalmente se classificam como "advérbios". Podemos atribuir a esse termo a função de **complemento do SAdj** (CSA).

Ao outro termo, *satisfeito,* atribuiremos a função de **núcleo do SAdj** (NSA). O NSA é preenchido por palavras únicas (e não seqüências de preposição + SN), tradicionalmente classificadas como "adjetivos" ou "substantivos". O NSA aparece logo antes do CSA ou, então, no final do sintagma, caso não haja CSA. Mas há, além disso, a possibilidade de ocorrência de um outro elemento logo após o NSA (e antes do CSA); é o caso de

(78) ... satisfeito demais com a vida.

Digamos que *demais* tem a função de **intensificador-1** (Int-1); veremos mais adiante por que "numerar" essa função.

Até o momento, as coisas não são excessivamente complexas. Poderíamos definir um SAdj máximo, ilustrado pelo exemplo (74); o CSA seria definido como ocorrendo em último lugar, o Int-1 como ocorrendo em penúltimo lugar, e o NSA como ocorrendo em antepenúltimo lugar no SAdj máximo.

Os problemas aparecem quando se consideram certos elementos que ocorrem antes do NSA. Vejamos os exemplos

(79) ... *muito* satisfeito com a vida.

(80) ... *sempre muito* satisfeito com a vida.

(81) ... *realmente sempre muito* satisfeito com a vida.

Os elementos em itálico nesses exemplos terão de ser incluídos no SAdj. E não é nada claro se se trata de novas funções ou da repetição da mesma função (no caso, intensificadores repetidos). No momento em que escrevo estas linhas, parece-me que é levemente preferível analisá-

los como desempenhando funções distintas; mas preciso deixar claro que as razões para isso não são fortes, de modo que a solução oferecida é necessariamente provisória.

A razão principal para distinguir aí várias funções é que não é possível inverter livremente a ordem dos elementos em um sintagma como (81):

(82) * ... sempre realmente muito satisfeito com a vida.
(83) * ... muito realmente sempre satisfeito com a vida.
(84) * ... realmente muito sempre satisfeito com a vida.

Isso parece indicar uma diferença de comportamento sintático e, portanto, uma diferença de função. Por isso, estabeleceremos três novas funções, todas chamadas **intensificadores**, e continuaremos a numeração iniciada com o Int-1, ou seja:

(85) ... realmente sempre muito satisfeito com a vida.

Int-4 Int-3 Int-2

Dessa maneira, a estrutura do SAdj (nesta concepção provisória) é a seguinte:

[Int-4 Int-3 Int-2 NSA Int-1 CSA]

A distribuição dos itens léxicos que podem desempenhar as quatro funções denominadas "intensificadores" é relativamente complexa. Existem, no mínimo, quatro classes de itens: os que só podem ser Int-1 (*demais*); os que só podem ser Int-2 (*muito*); os que podem ser Int-3 ou Int-1 (*sempre*); e os que podem ser Int-4 ou Int-1 (*realmente*). Os itens destas duas últimas classes parecem ser relativamente numerosos; já *demais* e *muito* são, aparentemente, representantes únicos de suas classes — ou seja, *demais* é a única palavra que **só** pode ser Int-1, e *muito* a única que **só** pode ser Int-2.

Voltando ao CSA, é necessário observar que pode haver mais de um por sintagma; assim, podemos ter

(86) ... batido [a máquina] [com cuidado].

onde há dois CSAs. Outros exemplos de CSA repetido são

(87) ... traduzido do inglês com perfeição.
(88) ... espalhado pelo chão desordenadamente.

Note-se que, em todos esses casos, a inversão da ordem dos dois termos dá resultado aceitável, o que é indicação de que se trata da mesma função repetida, e não de duas funções diferentes.

4.5.1.2. Fatos e perguntas

Vamos resumir o que sabemos acerca da estrutura interna do SAdj. Em primeiro lugar, parece que o SAdj máximo compreende seis termos, a saber,

[Int-4 Int-3 Int-2 NSA Int-1 CSA].

Vimos também que o CSA é preenchido por uma seqüência de preposição + SN ou, então, por uma palavra formada com o sufixo *-mente.*

Quanto aos diversos intensificadores, parece que são preenchidos por um número relativamente pequeno de itens: *muito, demais, realmente, sempre, já, agora* e alguns outros (veja-se, porém, abaixo uma dúvida a esse respeito). Desse modo, pode-se dizer que há certo grau de correlação entre classes de unidades (palavras ou sintagmas) e as funções que elas podem desempenhar no SAdj. Isso, no momento, é pouco mais que um palpite, uma sugestão de pesquisa.

Outra coisa que se pode afirmar é que pelo menos um tipo de restrição de compatibilidade funciona dentro do SAdj: a ocorrência de qualquer dos intensificadores só é possível na presença de um núcleo do SAdj. Fora essa restrição, a ocorrência dos termos é livre. Assim, um SAdj pode ser formado por um NSA ou por NSA + CSA, com ou sem intensificador(es); ou por um CSA, desta vez sem intensificador. Por exemplo:

(89) ... um homem [satisfeito] (NSA)

(90) ... um homem [satisfeito com a vida] (NSA + CSA)

(91) ... um homem [de Manaus] [CSA]

(92) ... um homem [muito satisfeito com a vida] (Int-2 + NSA + CSA)

(93) ... * um homem [muito de Manaus] (Int-2 + CSA)

As restrições de compatibilidade nos fornecem uma indicação para decidir a análise de casos como

(94) ... um amigo *de meu pai.*

O termo grifado parece um CSA, porque é composto de preposição mais SN; mas poderia, é claro, ser um exemplo de NSA, que nesse caso teria essa composição. Duas razões, entretanto, nos levam a preferir a análise de *de meu pai* como CSA: primeiro, os casos claros, como o de

(95) ... um homem satisfeito com a vida.

onde *com a vida* é CSA (e não pode trocar de posição com *satisfeito,* que é o NSA).

A segunda razão é o fato de que, em (94), não se pode inserir um intensificador:

(96) ... * um amigo muito de meu pai.

Isso ficará automaticamente explicado se *de meu pai* for um CSA, pois sabemos que o Int só poderá ocorrer se houver NSA no sintagma.

Outra observação interessante, derivada da análise de Lemle (1984), é que a função que aqui denominamos CSA provavelmente engloba o tradicional "agente da passiva". Com efeito, o comportamento sintático parece ser semelhante, se não idêntico, embora a bem dizer falte aqui (como em geral no caso do SAdj) um levantamento suficientemente sistemático e amplo. De

qualquer forma, analisaremos o elemento grifado em

(97) Compramos um biombo pintado *por Guignard.*

como CSA. Isso quer dizer que *pintado por Guignard* é um SAdj, do qual *pintado* é o núcleo. A análise de Lemle implica analisar orações passivas como

(98) Este biombo foi pintado por Guignard.

como construções compostas de sujeito + verbo (no caso, *ser* ou *estar*) + complemento do predicado (segundo a análise proposta nesta *Gramática*; tradicionalmente, "predicativo do sujeito"). Isto faz com que as passivas sejam paralelas, em certo nível de análise, a frases como

(99) Este biombo é importado.

O relacionamento semântico das passivas com "suas" ativas se processa através de regras especiais.

Finalmente, sinto-me obrigado a apontar desde já certos casos mal compreendidos, que poderão constituir problema para a análise do SAdj. Esses casos lançam dúvidas sobre a correlação entre classe e função que se observa em certas áreas do SAdj.

Uma dessas correlações seria a de que uma seqüência de preposição + SN, dentro de um SAdj, só poderia funcionar como CSA. Mas há duas classes de possíveis exceções; e, se se verificar que são realmente casos de outras funções desempenhadas por tais seqüências, poderemos vir a negar aquela correlação. Por exemplo, temos sintagmas como

(100) ... um funcionário *com freqüência bêbado.*

A porção em negrito funciona claramente como modificador e é, portanto, um SAdj. Mas é difícil negar que o núcleo é *bêbado,* o que nos obrigaria a analisar *com freqüência* como intensificador — muito embora se trate de uma seqüência de preposição + SN. A outra possibilidade seria analisá-la como CSA, mas isso nos obrigaria a admitir que o CSA pode ocorrer antes do NSA, e não há evidências independentes em favor disso. Temos aqui, portanto, um problema a resolver.

Outros casos parecem indicar que a seqüência preposição + SN pode ocorrer, ainda que raramente, como NSA. Nesses casos, ocorre um SAdj composto de intensificador + seqüência de preposição + SN:

(101) ... um empregado *muito de confiança.*

Como sabemos que o Int não pode ocorrer na ausência de um NSA, parece inevitável analisar *de confiança* como núcleo. Poderíamos salvar a análise sem maiores modificações argumentando que não se trata propriamente de sintagmas com preposição, mas de expressões fixas, como *sem vergonha* ou *para-casa*

(como em *faça o seu para-casa*). Ainda aqui, entretanto, fazem-se necessários estudos mais detalhados.

A análise acima é produto de um exame bastante apressado dos fatos; é possível que um estudo mais cuidadoso venha a modificar sensivelmente o quadro apresentado. Até que se realize tal estudo, contentemo-nos com esse precário esboço da estrutura interna do sintagma adjetivo.

4.6. O SINTAGMA ADVERBIAL

Dos chamados "sintagmas adverbiais", há pouco o que dizer no momento. Eles constituiriam a classe dos constituintes que ocupam funções "adverbiais" na oração, como, por exemplo, em

(102) Terminamos a pintura *em poucas horas.*

O constituinte *em poucas horas* [atributo em (102)] seria um sintagma adverbial.

Mas a simplicidade do conceito é só aparente, porque as funções ditas "adverbiais" são bastante diferentes entre si. As funções tradicionalmente chamadas "adverbiais" corresponderiam, no nível da oração, às funções de atributo, adjunto adverbial, adjunto oracional, negação verbal e muitos casos de adjunto circunstancial; além disso, também corresponderiam, em nível suboracional, ao intensificador e ao complemento do SAdj: como se vê, um grupo bastante grande e heterogêneo de funções.

O eventual estabelecimento de uma categoria de "sintagmas adverbiais" dependeria de se mostrar que todas essas funções, ou a maioria delas, são preenchidas por uma classe única de sintagmas — o que é certamente falso. O caso mais evidente é o da negação verbal, ocupada por constituintes específicos, que não podem desempenhar nenhuma das outras funções.

Nos outros casos, observa-se muita polivalência funcional entre os sintagmas, mas há também sintagmas que só podem ter uma função; os fatos simplesmente não foram suficientemente destrinçados.

Por exemplo, em (103), temos um adjunto oracional:

(103) Com franqueza, acho que eles nos enganaram.

O sintagma *com franqueza* pode também ocorrer como adjunto circunstancial em

(104) Ele se expressou com franqueza.

Mas há sintagmas que podem ocorrer só em uma das duas funções. Em

(105) Ele se expressou com bom gosto.

temos um sintagma, *com bom gosto*, que é AC, mas não poderia ser AO. Ao contrário, o AO de

(106) Indubitavelmente, eles nos enganaram.

não poderia ocorrer como AC.

Como não se fez ainda um estudo das funções possíveis de tais sintagmas com base em um conjunto confiável de funções sintáticas, torna-se difícil no momento definir não só o "sintagma adverbial" (que, de qualquer modo, não deve ser uma categoria única), mas ainda um conjunto de funções que ocupem o lugar do "sintagma adverbial". O mais que se pode fazer é apontar algumas características que parecem ocorrer com freqüência nesses casos.

A estrutura interna desses sintagmas, segundo Huddleston, seria relativamente simples. Ele comenta apenas que

> Poucos advérbios aceitam complementos (como, por exemplo, *independently* aceita um sintagma preposicionado com *of*: *independently of such factors*), e a modificação é geralmente do tipo gradação, à maneira da que foi discutida no caso dos sintagmas adjetivos. [Huddleston, 1984, p. 335]

Pode-se dizer um pouco mais: esses sintagmas podem ser formados de preposição + SN, ou de um item "adverbial" acompanhado ou não de complemento. Em especial, ocorre um complemento algo semelhante ao intensificador que distinguimos no SAdj:

(107) Ele se expressou *muito francamente*.

Existe, sem dúvida, uma relação estrutural bastante estreita entre o SAdj e certos tipos de "sintagmas adverbiais": para ver isso, basta comparar os sintagmas *muito francamente* e *muito franco*. Observa-se tradicionalmente que é como se fossem duas variantes funcionais de uma mesma estrutura.

No entanto, a falta de uma definição (ou melhor, um conjunto de definições) que delimite com alguma clareza as entidades de que estamos falando impede qualquer tentativa séria de sistematização. Prefiro, por isso, limitar-me às observações gerais dadas acima, aguardando que se elabore uma taxonomia rigorosa desses sintagmas.

4.7. COMPLEMENTOS X ADJUNTOS

Encontra-se com freqüência nas gramáticas tradicionais uma classificação das funções sintáticas segundo a qual algumas seriam "essenciais", outras "integrantes", outras "acessórias"; ou, ainda, algumas seriam "complementos" e outras "adjuntos". Vou fazer a seguir um breve exame da relevância dessas classificações.

Tomemos primeiro a distinção entre termos (ou funções) "essenciais", "integrantes" e "acessórios". Segundo alguns, a diferença reside na importância relativa dos diversos termos (os mais importantes seriam os essenciais, e assim por diante). Como não se especifica o que vem a ser "importância" nesse contexto, fi-

ca difícil aceitar, e mesmo criticar, essas noções, quando assim definidas.

Uma outra concepção desses três tipos de funções é a de que os termos "essenciais" teriam que estar presentes em qualquer oração; os "integrantes" seriam, em certos casos, "exigidos" por algum outro elemento da oração; e os "acessórios" não seriam em nada importantes para o sentido da oração (como se vê, a definição tende a ser em parte semântica). A aceitarmos essa definição, o único termo "essencial" seria o predicado.

Já a distinção entre "integrantes" e "acessórios" teria como base a possibilidade ou não de serem exigidos por outro termo da oração; nessa acepção, os termos "integrantes" seriam os que alguns denominam "complementos", e os "acessórios" seriam os "adjuntos". Assim, poder-se-ia dizer que o objeto direto é um "complemento" (ou que é um "termo integrante"), porque existem verbos, como *fazer* e *achar*, que exigem sua presença nas orações em que aparecem; já o "adjunto adverbial" seria um "adjunto" (um "termo acessório"), porque nenhum verbo, substantivo, adjetivo etc. requer sua presença.

Nesse caso, os "termos integrantes" ou "complementos" seriam para nós apenas quatro: o **objeto direto**, o **complemento do predicado**, o **predicativo** e o **adjunto circunstancial**; isso porque são essas as únicas funções exigidas (ou recusadas) por certos verbos. Todas as outras funções são de ocorrência livre com todos os verbos (ver a seção 6.2., onde se estuda em detalhe o fenômeno da transitividade).

Como o leitor pode notar, não incorporei a oposição entre "adjuntos" e "complementos" na nomenclatura adotada. Assim, como acabo de apontar, o adjunto circunstancial seria um complemento; no entanto, não me pareceu necessário manter essa denominação, já que ela apenas repetiria a distinção, estabelecida na seção 6.2., entre funções relevantes e irrelevantes para efeito de formulação das transitividades. Em resumo, as denominações "termo essencial", "integrante" e "acessório" não são utilizadas na presente análise; e os termos "complemento" e "adjunto", embora utilizados aqui, não se distinguem em termos de sua relevância para a formulação das transitividades.

4.8. O APOSTO E OS PARENTÉTICOS

4.8.1. Os parentéticos

Antes de encerrar este capítulo, vou fazer algumas considerações sobre os **parentéticos**: elementos que podem posicionar-se livremente entre os constituintes oracionais e que na escrita são sempre separados por vírgula.

Os parentéticos ainda não foram estudados com o cuidado que sem dúvida merecem; mas já se pode vislumbrar o suficiente para sugerir que sua análise pode vir a ser importante para uma melhor compreensão de certos termos de comportamento algo obscuro, como o "aposto" da gramática tradicional. Além disso, é possível que a análise dos parentéticos nos obrigue a uma reconsideração do adjunto oracional, estudado no capítulo 3, e mesmo de outros termos de nível oracional.

Um exemplo bem típico de ocorrência de parentético são as frases abaixo:

(108) a. Creio eu, Dorival dispensou o sócio.
b. Dorival, creio eu, dispensou o sócio.
c. Dorival dispensou, creio eu, o sócio.
d. Dorival dispensou o sócio, creio eu.

Em geral se sustenta que parentéticos como *creio eu* não têm relação sintática nem semântica com a oração onde se inserem; quanto a isso, terei algo a dizer mais adiante.

A estrutura interna dos parentéticos sugere que eles realmente não são termos da oração; por exemplo, em (108), o parentético tem a estrutura de uma oração independente, isto é, não apresenta nenhuma das marcas habituais de subordinação listadas na seção 5.2.3.: conjunção, ou desinência de infinitivo ou gerúndio etc. Também não apresentam as marcas de coordenação listadas em 5.2.4. As frases de (108) parecem antes formadas de duas orações independentes, simplesmente justapostas.

Essa justaposição não se verifica somente com orações; pode ocorrer, ao que tudo indica, também com termos da oração, que são repetidos e separados por vírgula:

(109) Mandei um cartão ao Zé, ao meu velho companheiro de lutas.

(110) Os deputados dão, oferecem de graça, empregos na Assembléia.

Em (109) temos a repetição justaposta de um atributo, e em (110) de um predicado.

Digamos que se trata do mesmo fenômeno nos três casos; se for assim, então o chamado "aposto" será mais um exemplo do mesmo processo:

(111) Simone, irmã do Carlinhos, ganhou um carro novo.

Não vejo razão para analisar (111) diferentemente de (108), (109) e (110).

Nesse caso, poderemos definir os parentéticos como elementos que sintaticamente repetem a oração ou um de seus termos e se justapõem ao elemento repetido, separando-se dele por vírgula. Não podem, portanto, ser considerados um termo da ora-

ção, mas antes manifestações de um processo mais geral de repetição, que vale para muitos dos termos da oração ou dos sintagmas.

Entendo "repetição sintática" no sentido de que se repete a categoria, embora seu preenchimento e estrutura interna possam ser muito diferentes. Assim, em (108), temos uma oração justaposta a outra oração; mas cada uma tem uma estrutura e preenchimento léxico próprios. Em (109) justapõem-se dois sintagmas com função de adjunto circunstancial, em (110) dois sintagmas com função de predicado, em (111) dois sintagmas nominais. Não é necessário, portanto, especificar uma função separada para o "aposto" da gramática tradicional: trata-se simplesmente de um caso especial de parentético — o caso em que o elemento repetido é um SN.

4.8.2. A considerar: e o AO?

Essa análise sugere que o **adjunto oracional** (AO), definido como um termo da oração no capítulo 3, poderia ser mais um caso de parentético. Com efeito, o AO tem comportamento sintático algo semelhante ao dos parentéticos:

(112) a. Felizmente, Dorival dipensou o sócio.
b. Dorival, felizmente, dispensou o sócio.
c. Dorival dispensou, felizmente, o sócio.
d. Dorival dispensou o sócio, felizmente.

Só não me decidi a identificar o AO aos parentéticos na presente análise porque o AO, ao contrário dos parentéticos, não repete um termo de idêntica classe dentro da oração. Por exemplo, em (112), não há nenhum sintagma adverbial que se possa dizer repetido por *infelizmente.* Por isso, manteremos por ora a análise do AO como um não-parentético. Mas as semelhanças entre os dois tipos de elementos são grandes, e pode valer a pena rever essa análise no futuro.

4.8.3. Integração dos parentéticos na oração

Finalmente, uma rápida nota sobre a integração sintática e semântica dos parentéticos na oração.

O parentético, no que pese sua independência de posicionamento na oração, vincula-se a esta pela concordância nominal (quando aplicável) e pela interpretação semântica. Assim, em

(113) Irritado, Dorival dispensou a secretária.

o *irritado* concorda com *Dorival* e se refere semanticamente a ele: o irrita-

do é claramente Dorival, não a secretária. O oposto se verifica em

(114) Dorival dispensou a secretária, irritada.

O enunciado (114) me parece algo menos natural do que (113), certamente devido a uma tendência muito generalizada de se vincular elementos justapostos ao tópico da sentença.

Mesmo quando o parentético é de tipo mais independente, como os que tomam como escopo a oração inteira, é possível detectar vinculação semântica com certas partes da oração. Assim, percebo uma diferença de significado entre

(112) a. Felizmente, Dorival dispensou o sócio.

e

(112) c. Dorival dispensou, felizmente, o sócio.

Em (112c), o falante parece exprimir sua aprovação quanto à pessoa dispensada: Dorival poderia ter dispensado, digamos, a secretária ou o sócio, e o falante deixa claro que a decisão de dispensar o sócio foi acertada. Em (112a), não vejo possibilidade de uma interpretação nessa linha.

5 A oração complexa

5.1. RECURSIVIDADE

No capítulo 3, estudamos a estrutura da oração simples, isto é, da oração que não contém, ela própria, outras orações. Agora vamos passar ao estudo da oração complexa, definida como uma oração que contém dentro de seus limites pelo menos uma outra oração.

Um exemplo de oração complexa é

(1) Titia disse que nós desarrumamos a casa.

Essa frase pode ser analisada como contendo um sujeito, *titia*, um núcleo do predicado, *disse*, e um objeto direto, *que nós desarrumamos a casa*. O que nos interessa no momento é o objeto direto, pois ele contém uma oração, *nós desarrumamos a casa*; essa oração, por sua vez, tem sujeito (*nós*), NdP (*desarrumamos*) e OD (*a casa*). Utilizando colchetes como sinal de limite de oração, podemos representar a frase da seguinte maneira:

[Titia disse que (nós desarrumamos a casa)]

Trata-se, como se vê, de uma estrutura (uma oração) dentro de outra estrutura da mesma classe (outra oração). Essa propriedade, comum a todas as línguas, de colocar estruturas dentro de outras estruturas da mesma classe, se chama **recursividade.**

A recursividade é uma das propriedades mais importantes das línguas humanas, pois é principalmente ela que permite aos falantes produzir um número potencialmente ilimitado de sentenças. A recursividade distingue as línguas humanas das linguagens de certos animais, compostas estas de um número fixo de sinais e que portanto só podem transmitir um número limitado de mensagens. (Isso não vale para **todas** as linguagens animais; algumas, como a dança das abelhas, têm possibilidades ilimitadas de criação de mensagens.)

Em português podemos ter não apenas orações dentro de orações [como no exemplo (1)], mas SNs dentro de SNs, como em

(2) Os dedos dos pés de Genita

onde *os dedos dos pés de Genita* é um SN, pois pode ser sujeito (OD...) de uma oração; igualmente, *os pés de Genita* é um SN; e *Genita* também é um SN. Podemos ainda colocar sintagmas adjetivos dentro de sintagmas adjetivos, e assim por diante.

5.2. ORAÇÕES DENTRO DE ORAÇÕES

Nesta seção, interessa-nos o caso especial de orações dentro de orações. Voltemos ao exemplo

(1) Titia disse que nós desarrumamos a casa.

A sequência *nós desarrumamos a casa* é uma oração, porque tem a estrutura típica de uma oração. Se aplicarmos a essa seqüência, isolada, os testes vistos na seção 3.2. para a atribuição de traços e funções, veremos que *nós* é sujeito porque é marcado [+CV], e *a casa* é OD porque é marcado [–CV, + Ant, +Q, –PA...].

Chamamos, então, (1) de **oração complexa** (em oposição às **orações simples** vistas até agora).

5.2.1. Critério de contagem de orações

Uma oração complexa pode ser identificada porque repete duas ou mais vezes a estrutura típica de uma oração. A estrutura típica da oração varia bastante — algumas orações têm sujeito, outras não têm, algumas têm OD, outras não etc. — mas em geral não é difícil identificar uma oração, porque pelo menos algumas marcas bem características estão geralmente presentes. Ora, sabemos que cada oração tem um (e só um) NdP; e não pode ter mais de um sujeito, um OD, um CP e um predicativo. Com base nessas observações, é possível examinar uma oração e decidir se é complexa.

5.2.1.1. Casos de segmentação clara

Casos bem favoráveis são os de

(1) Titia disse que nós desarrumamos a casa.

(3) Titia fez a salada e mamãe fritou os pastéis.

Em (1), há claramente dois NdPs, *disse* e *desarrumamos*; não há meio de combiná-los em um predicado complexo (ver o conceito de “predicado complexo” na seção 3.2.2.2.), pois não estão contíguos e, ademais, estão ambos em forma pessoal. Há também dois sujeitos, um

para cada NdP. Só isso já é o bastante para que concluamos que há aí duas orações. Um raciocínio análogo se aplica a (3), onde há também duas orações (mais a oração que as engloba, como veremos na seção 5.2.2.3.).

Além das razões aventadas acima, podemos acrescentar que seqüências como *titia fez a salada* ou *nós desarrumamos a casa* podem ocorrer como frases independentes, o que reforça seu caráter oracional. No entanto, esse critério falha com muitas orações, em especial com certas subordinadas (com as de gerúndio, por exemplo), de modo que esse critério não pode ser o decisivo. Estabelecerei, pois, que uma oração é complexa quando um ou mais de seus termos têm a estrutura interna típica de uma oração.

O critério adotado, que se aplica tão facilmente a casos como (1) e (3), apresenta alguma dificuldade com outras estruturas. Vou agora examinar os diversos tipos de estruturas que poderiam ser consideradas orações complexas, para ver quais delas efetivamente o são.

Em uma ordem aproximada de "oracionalidade decrescente", essas estruturas são:

(a) construções com *que* + verbo no indicativo;

(b) construções com *que* + verbo no subjuntivo;

(c) construções de infinitivo;

(d) construções relativas;

(e) construções de gerúndio;

(f) construções de particípio.

E a pergunta é: quais dessas construções devem ser consideradas orações (de modo que a oração em que ocorram como termo seja considerada uma oração complexa)?

5.2.1.2. Com verbos no subjuntivo

O caso (a) corresponde aos exemplos já examinados, (1) e (3). Aqui a decisão é fácil porque a oração subordinada tem estrutura idêntica à de uma oração que apareça sozinha (oração **absoluta**, na terminologia tradicional). Mas já com o caso (b) isso não acontece:

(4) O governo quer que o povo aperte o cinto.

A seqüência *o povo aperte o cinto* não aparece normalmente como oração absoluta; será, apesar disso, uma oração?

A resposta deve ser positiva, porque, exceto a forma particular do verbo, os demais sinais da existência de uma oração estão presentes: há um sujeito, *o povo*, com o qual o NdP claramente concorda; e não é difícil mostrar, através da transitividade de *apertar*, que *o cinto* é o OD. Logo, trata-se de uma oração, embora sua forma seja "especializada", só podendo ocorrer como subordina-

da. (Há, na verdade, alguns casos em que uma oração com verbo no modo subjuntivo pode ocorrer como absoluta; mas nem sempre isso é possível, e de qualquer modo a observação não é importante aqui.)

5.2.1.3. Orações "reduzidas"

Passemos agora ao caso (c), construções de infinitivo. Quando se trata de infinitivo pessoal, como em

(5) Foi uma beleza vocês chegarem.

o mesmo raciocínio aplicado a (4) vai funcionar, mostrando que *vocês chegarem* é uma oração. Já no caso de

(6) Os juízes decidiram considerar ilegal o imposto.

as coisas não são tão claras. Na seqüência *considerar ilegal o imposto*, não há um verbo que concorde, e portanto não há sujeito.

Ainda assim é possível mostrar que *considerar ilegal o imposto* é uma oração separada. Para verificar isso, relembremos a discussão da seção 3.2.2.2., em que introduzi a noção de "predicado complexo". Naquela seção, foi proposto um critério para verificar se uma seqüência de verbo em forma finita + infinitivo (gerúndio, particípio) é um predicado complexo (e, portanto, uma única oração) ou dois predicados (e duas orações). O critério se baseia nos traços de transitividade, partindo do princípio de que cada oração tem um predicado e, portanto, um conjunto de traços de transitividade. Assim, na frase

(7) Daniel vai correr o mundo.

a presença (opcional) de um OD, assim como a impossibilidade de se acrescentar um predicativo são atribuíveis exclusivamente aos traços de transitividade do verbo *correr*; o verbo *vai* não tem influência na transitividade. Tanto é assim que as condições de transitividade de frases com *vai correr* são em tudo idênticas às de frases com a forma *correu*, por exemplo. A partir de observações como essa, chegamos à conclusão geral de que *ir* + infinitivo forma um predicado complexo e que (7) contém apenas uma oração.

Nos casos em que se pode perceber **dois** conjuntos de traços de transitividade, conseqüentemente, deverá haver duas orações. Por exemplo, na frase

(6) Os juízes decidiram considerar ilegal o imposto.

é possível mostrar que os dois verbos impõem cada um suas próprias exigências quanto aos complementos possíveis. Assim, *ilegal* é um predicativo; mas o verbo *decidir* não admite predicativo, como se pode ver em

(8) * Os juízes decidiram ilegal o imposto.

Portanto, a presença do predicativo em (6) deve ter sido autoriza-

da pela transitividade de *considerar*; isso significa que *considerar* tem, por assim dizer, sua área própria de jurisdição, ou seja, uma oração separada. Conclui-se que (6) tem duas orações.

À primeira vista, esse processo de verificação parece algo complicado; mas é claro que, na prática, já sabemos que os verbos que entram na composição de predicados complexos são muito poucos, a saber:

ir (+ infinitivo);

ter, haver (+ particípio);

estar, vir, ir, andar (+ gerúndio);

ser (*estar*) (+ particípio)
[= construção passiva].

Além disso, dessas oito combinações possíveis, somente três (*vir, ir* e *andar* + gerúndio) têm chances de ocorrer fora de predicados complexos; e uma (a construção passiva) é com freqüência homônima de uma seqüência de *ser* + adjetivo. As outras quatro combinações dispensam testagem; e, naturalmente, qualquer combinação que não esteja nessa lista de oito (como a combinação *decidiram considerar*) igualmente dispensa testagem. Portanto, na prática, a verificação é fácil; mas é importante ter em mente a argumentação acima, que é a justificação teórica da análise adotada.

Voltando à nossa lista de construções, até agora verificamos que: (a) as construções de indicativo e de subjuntivo sempre constituem orações separadas; (b) as construções de infinitivo constituem orações separadas, exceto quando integram predicados complexos.

Acrescentarei que a situação das construções de gerúndio é idêntica à das construções de infinitivo. Assim, temos uma oração em

(9) Daniel está vendendo cimento.

e duas orações em

(10) Daniel trabalha vendendo cimento.

Já a situação das construções de particípio é diferente. Veremos que o particípio **nunca** forma uma oração separada; na verdade, quando ocorre fora de um predicado complexo, o particípio tem as características de um adjetivo. Nesse caso, o particípio não apenas concorda em gênero e número (o que não acontece quando faz parte do predicado complexo), mas seus complementos têm a forma de complementos de adjetivo, e não de complementos de verbo.

Assim, em uma frase como

(11) Esses costumes vieram da Europa.

o constituinte *da Europa* se analisa como **adjunto circunstancial**, isto é, [–CV, +Ant, –Q, –CN, +Cl, –PA]. Mas no sintagma

(12) Costumes trazidos da Europa

o mesmo constituinte, embora de forma e significado idênticos, tem claramente a função de comple-

mento do sintagma adjetivo. Por exemplo, não pode ser anteposto nem clivado:

(13) * Da Europa, ele adotou costumes trazidos.

(14) * Foi da Europa que ele adotou costumes trazidos.

Em uma palavra, *da Europa* em *trazidos da Europa* não é um constituinte de nível oracional. Por isso, analisarei uma frase como

(15) Ele adotou costumes trazidos da Europa.

como tendo apenas uma oração; e *trazidos da Europa* é um modificador dentro do SN *costumes trazidos da Europa.* Esse modificador não é oracional, ou seja, não contém uma oração.

5.2.1.4. Conclusão

Devemos concluir, pois, que existem em português dois tipos de "orações reduzidas", isto é, cujo NdP não se flexiona em número e pessoa: as de infinitivo e as de gerúndio. As chamadas "reduzidas de particípio" não são orações, mas sintagmas adjetivos bastante regulares.

Os testes ilustrados acima devem ser suficientes para decidir, na maioria dos casos, quantas orações contém uma sentença. É de esperar que haja casos problemáticos, mas devem ser uma minoria. Passemos agora ao estudo dos modos de organização dos constituintes oracionais (que contêm oração) dentro da oração complexa.

5.2.2. Subordinação e coordenação

Tradicionalmente, distinguem-se duas maneiras básicas de inserir constituintes dentro de outros constituintes: a subordinação e a coordenação. Essa dicotomia representa uma simplificação e está longe de dar conta de todos os fatos, mas servirá como ponto de partida para a exposição, de modo que vou começar conceituando esses dois processos.

Em um exemplo como

(1) Titia disse que nós desarrumamos a casa.

uma das orações, *nós desarrumamos a casa,* está inserida dentro de um dos termos da outra oração (esta é *titia disse que nós desarrumamos a casa,* ou seja, a íntegra da sentença): está dentro do objeto direto: *que nós desarrumamos a casa* (veremos em 5.2.2.2. por que não considerar *que* como parte da oração menor).

Uma razão para se adotar essa análise deriva da transitividade do verbo *dizer.* Com efeito, esse verbo parece exigir objeto direto:

(16) Titia disse algumas asneiras.

(17) * Titia disse.

Ora, o constituinte *que nós desarrumamos a casa,* além de estar no lugar habitual do objeto direto, logo depois do NdP, parece satisfazer a transitividade de *dizer,* de modo que a frase (1) fica aceitável. Se esse constituinte não fosse objeto direto, teríamos de explicar por que (1) é bem formada, apesar de conter o verbo *dizer* sem objeto direto. Analisando *que nós desarrumamos a casa* como objeto direto, esse problema desaparece.

Além da transitividade do verbo, outros traços do constituinte *que nós desarrumamos a casa* sugerem que este tem a função de OD. Por exemplo, é possível antepô-lo e mesmo clivá-lo:

(18) Que nós desarrumamos a casa, titia disse.

(19) Foi que nós desarrumamos a casa que titia disse.

Ou seja, podemos dizer que a oração *que nós desarrumamos a casa* é marcada [+ Ant, +Cl], reforçando assim as indicações de que se trata de um objeto direto. Nesse caso, (1) compreende duas orações, sendo que uma delas (*nós desarrumamos a casa*) está **dentro** da outra (*titia disse que nós desarrumamos a casa*), no sentido de que faz parte de um de seus termos.

Em casos como esse, em que uma oração faz parte de um termo de outra, falamos de subordinação; assim, a oração menor, *nós desarrumamos a casa,* é a oração subordinada; e a maior, *titia disse que nós desarrumamos a casa,* se denomina oração principal. A subordinação é um dos dois processos principais de montagem de orações complexas.

Já na sentença

(3) Titia fez a salada e mamãe fritou os pastéis.

temos igualmente pelo menos duas orações: uma delas é *titia fez a salada,* e a outra é *mamãe fritou os pastéis.* No entanto, não se pode dizer que alguma delas esteja "dentro" da outra, no sentido de que faz parte de um de seus termos. Tanto é assim que ambos os verbos, *fez* na primeira e *fritou* na segunda, têm todos os complementos exigidos preenchidos por elementos não-oracionais: na primeira, o sujeito é *titia,* o NdP é *fez,* o OD é *a salada.* Nesse caso, diz-se que a sentença é composta por coordenação; e as duas orações são coordenadas, isto é, sintaticamente equivalentes. A coordenação é, pois, o segundo dos processos principais de montagem de orações complexas. Naturalmente, nenhuma das orações coordenadas pode ser chamada de "principal", já que "principal" se define em termos da presença de subordinadas.

É só nos casos de subordinação que se verifica a coincidência parcial das duas orações, de modo que, quando citamos as duas orações, uma parte deverá ser citada duas vezes: a principal de (1) é *titia disse que nós desarrumamos a casa,* e a subordinada é *nós desarrumamos a casa.* Nos

casos de coordenação, as duas orações são totalmente separadas, e nenhum elemento é parte de ambas: as duas coordenadas de (3) são *titia fez a salada* e *mamãe fritou os pastéis.*

É também só nos casos de subordinação que, freqüentemente, uma das orações tem uma forma especial que nunca, ou só muito raramente, ocorre isolada: orações de subjuntivo, gerúndio e infinitivo são sempre subordinadas. Na coordenação sem subordinação (porque os dois processos podem coexistir na mesma sentença, ver a seção 5.2.2.4.), as orações têm todas a mesma forma que teriam se ocorressem isoladas.

Toda a exposição acima tratou da coordenação e subordinação de **orações** umas às outras. No entanto, essas noções se aplicam igualmente a outras formas sintáticas. Assim, em um sintagma como *Pedro e Simão,* podemos dizer que há dois SNs, *Pedro* e *Simão,* coordenados um ao outro e formando, em seu conjunto, um SN maior, *Pedro e Simão.* Isso pode acontecer com verbos, como em *viu e venceu,* com advérbios, como em *demorada e pacientemente* e assim por diante.

Igualmente, é possível subordinar constituintes não-oracionais uns aos outros. Por exemplo, em *a filha do vizinho,* temos um SN, *o vizinho,* fazendo parte de um SN maior, *a filha do vizinho;* trata-se, pois, de um típico caso de subordinação.

Vou agora passar a esclarecer três pontos que podem criar confusão, por se oporem de certa forma à prática tradicional da análise sintática.

5.2.2.1. Limites da oração principal

Um aspecto da análise que contraria a tradição tem a ver com os limites da oração principal. Eu disse acima que a oração principal de (1) é *titia disse que nós desarrumamos a casa.* Mas tradicionalmente se afirma que a principal é apenas *titia disse* — presumivelmente porque o resto do período, *que nós desarrumamos a casa,* teria de ser reservado para ser a subordinada.

Essa é a prática usualmente seguida na análise sintática. No entanto, vou mostrar que não só é inconveniente, como vai contra a própria doutrina tradicional tal como se explicita; ou seja, neste ponto (como em muitos outros) a prática e a teoria tradicionais estão em conflito.

Vejamos primeiro a definição tradicional de "oração subordinada"; todas as gramáticas trazem uma definição equivalente à seguinte:

> As orações sem autonomia gramatical, isto é, as orações que funcionam como termos [...] de outra oração chamam-se **subordinadas**. [Cunha & Cintra, 1985, p. 579]

Não creio que essa definição esteja correta sob todos os aspectos; mas sem dúvida capta o aspecto fundamental do fenômeno e, por isso, merece ser mantida na essência (haverá uma modificação secundária, como veremos logo adiante).

Voltando ao nosso exemplo,

(1) Titia disse que nós desarrumamos a casa.

vamos admitir a posição tradicional que segmenta *titia disse* como oração principal; tentarei mostrar que essa segmentação contradiz a definição de "subordinada".

Em primeiro lugar, observemos que **um termo de uma oração é parte dela**; assim, em

(20) Titia fez a salada.

a salada é um termo da oração e é parte da oração. Ninguém teria a idéia de dizer que a oração acima é apenas *titia fez*; essa seqüência não é mais que um pedaço de oração — uma oração da qual se extraiu o objeto direto.

Tudo isso parece insuportavelmente óbvio; mas deixa de ser observado quando se analisa (1). Se *que nós desarrumamos a casa* é um termo da principal (o OD), então *que nós desarrumamos a casa* deve fazer parte da principal. Se isso não acontecer, teremos que: (a) a principal, que então seria *titia disse*, **não teria objeto direto** (já que o OD precisa fazer parte da sua oração); ou então (b) o que chamamos "oração principal", *titia disse*, não seria na verdade uma oração, mas apenas um pedaço de oração (amputada do OD). A única saída seria postular um princípio segundo o qual o OD (assim como os demais termos) de uma oração só faz parte dela se não contiver uma oração; se contiver uma oração, o OD e os outros termos serão externos a essa oração. Esse princípio é tacitamente admitido na prática e é a única base para a afirmação de que *titia disse* é a oração principal de (1), mas introduz complicações enormes e desnecessárias em toda a concepção da sintaxe. No mínimo, teríamos que reconhecer que ele não nos compra vantagem nenhuma: seria introduzido na teoria exclusivamente a fim de salvar uma prática de análise estabelecida.

Prefiro, portanto, manter na essência a definição de "subordinada" vista acima e adequar a prática da análise a ela. A definição é satisfatória porque exprime a importante propriedade que têm as línguas de desenvolverem termos da oração como novas orações — um aspecto da recursividade (ver a seção 5.1.).

Assim, a oração principal de (1) é *titia disse que nós desarrumamos a casa*; e a subordinada é *nós desarrumamos a casa* (deixemos em suspenso, por ora, o estatuto da palavra *que*). Vejamos brevemente as vantagens que nos traz essa segmentação.

Primeiro, torna-se possível manter, sem sofismas, a afirmação de que uma oração subordinada é parte da principal: é evidente para qualquer um que a seqüência *nós desarrumamos a casa* é parte de *titia disse que nós desarrumamos a casa.* Depois, não se torna necessário complicar a teoria com princípios inconvenientes, como o que diz que um OD pode não fazer parte de sua oração; manteremos, em vez, o que o bom senso nos dita, ou seja, que os termos de uma oração fazem sempre parte dela. Evitamos também situações confusas, como dizer que a oração *titia disse* tem objeto direto (onde está ele?). Com isso, finalmente, adequamos a teoria (definição de "subordinada") à prática [segmentação de (1)], fazendo da prática uma aplicação da teoria — senão, para que seria preciso construir uma teoria?

Só com essa solução podemos manter sem incoerência afirmações como a de que as subordinadas são

> [...] termos da frase desenvolvidos em oração. [Kury, 1985, p. 71]

5.2.2.2. Limites da oração subordinada

Agora examinemos o estatuto da palavra *que*: faz ou não faz parte da oração subordinada? Vou defender a idéia de que não faz e que, portanto, a subordinada de (20) é apenas *nós desarrumamos a casa,* e não *que nós desarrumamos a casa*[1].

A razão principal é que a seqüência *nós desarrumamos a casa* tem a estrutura interna e a distribuição típicas de uma oração: tem sujeito, NdP, objeto direto; e poderia ocorrer sozinha no período — como as orações em geral. Já a seqüência *que nós desarrumamos a casa* não poderia ocorrer sozinha e tem um elemento (*que*) sem função conhecida dentro da oração simples.

A seqüência *que nós desarrumamos a casa,* na verdade, apresenta semelhanças funcionais com os sintagmas nominais: pode ser objeto direto [como em (20)] e pode ser sujeito, como em

(21) Que nós desarrumamos a casa é evidente.

Assim, a oração subordinada de (1) é *nós desarrumamos a casa*; e o elemento *que* é um dos recursos que a língua possui para encaixar uma oração dentro de um SN; em outras palavras, a gramática portuguesa estipula que a conjunção *que,* mais uma oração, forma um SN.

Note-se que, com isso, já não se pode dizer que a oração subordinada é objeto direto da principal. Ela faz parte do objeto direto, mas este é preenchido pelo SN composto de ***que* + oração**. E podemos

[1] Refiro-me aqui à conjunção *que,* não ao pronome relativo *que*: este pertence à oração subordinada. A respeito dos relativos, ver 5.2.3.2.

manter a afirmação geral de que objetos diretos (e sujeitos) são sempre SNs.

Voltando agora à definição tradicional de "subordinada", já temos condições de fazer nela a pequena modificação a que aludi acima. Uma oração subordinada não é necessariamente a que funciona como termo de outra oração; mas é aquela que é **parte de um termo de outra oração**. *Nós desarrumamos a casa* é uma subordinada porque faz parte de *que nós desarrumamos a casa,* que por sua vez é termo de outra oração. Em outros casos, a subordinada constitui, sozinha, um termo da principal — é o que acontece, por exemplo, com subordinadas de gerúndio ou infinitivo; mas isso nem sempre acontece, e é preciso incorporar essa informação à definição.

Antes de passar à coordenação, uma observação importante: o termo "principal" é relativo, ou seja, uma oração é principal em relação a sua subordinada; mas essa mesma oração pode ser subordinada em relação a uma terceira. Isso acontece quando, em uma sentença, encontramos uma oração que é parte de outra, sendo que esta é, por sua vez, parte de uma terceira:

(22) Minha irmã sonhou que eu lhe disse que papai ia viajar.

Nessa sentença, a oração *papai ia viajar* é subordinada a *eu lhe disse que papai ia viajar*; esta última é, portanto, a principal de *papai ia viajar.* Mas essa mesma principal é, por sua vez, subordinada a *minha irmã sonhou que eu lhe disse que papai ia viajar.* A estrutura pode ser representada como

[minha irmã sonhou que (eu lhe disse que {papai ia viajar})]

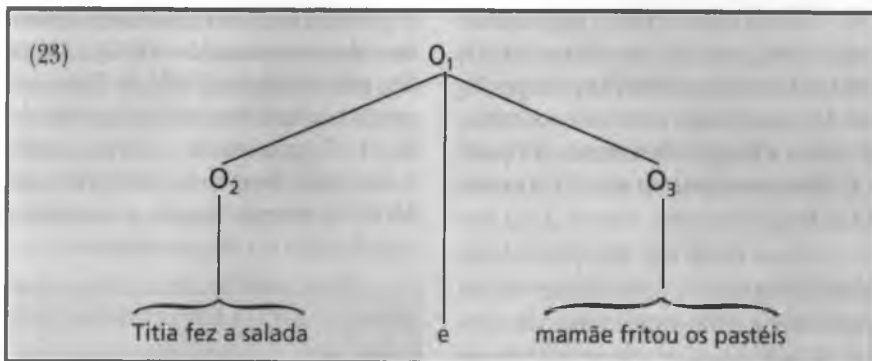
onde a oração entre parênteses, ( ), é principal da oração entre chaves, { }, e subordinada à oração entre colchetes, [ ].

5.2.2.3. Coordenadas: quantas orações?

Vamos agora apurar quantas orações devem ser distinguidas em

(3) Titia fez a salada e mamãe fritou os pastéis.

A questão é: temos aqui apenas duas orações? Ou três orações, sendo uma delas a íntegra da sentença, a saber, *titia fez a salada e mamãe fritou os pastéis?*

Vou adotar a posição de que (3) encerra três orações, a saber, (a) *titia fez a salada*; (b) *mamãe fritou os pastéis*; e (c) *titia fez a salada e mamãe fritou os pastéis.* Naturalmente, não há dúvida de que (a) e (b) são orações; vou mostrar por que (c) também deve ser considerada uma oração. A estrutura que proponho para (3) pode ser esquematizada assim:

A justificação direta dessa análise se baseia no seguinte argumento: se o conjunto das duas orações não formar uma oração, teremos de criar uma nova classe para acomodá-lo, e os membros dessa classe teriam pelo menos algumas das propriedades distribucionais das orações (por exemplo, ocorrer sozinhas no período). Por isso, essa nova classe seria suspeita de não passar de mais um rótulo para a classe "oração".

Por outro lado, há também um motivo indireto para adotar (23) como a melhor análise para (3): muitos exemplos sustentam o princípio de que, onde cabe uma classe **X**, cabe também a coordenação de duas (ou mais) ocorrências de **X**, isto é, **X e X**. E, nesses casos, a seqüência **X e X** se comporta gramaticalmente da mesma maneira que **X** sozinho, merecendo portanto o rótulo de **X**. Isso se verifica com o SN:

(24) a. Avistei hoje seu patrão.
b. Avistei hoje seu patrão e o doutor Márcio.

com o verbo:

(25) a. Vocês roubaram meu chocalho.
b. Vocês roubaram e estragaram meu chocalho.

com o adjetivo:

(26) a. Um simples auxiliar
b. Um simples e competente auxiliar

e assim por diante.

Em cada um desses casos, a junção dos dois elementos forma um novo sintagma, cuja função é a mesma que teria um dos elementos se ocorresse só: em (24b), *seu patrão e o doutor Márcio* é um objeto direto; em (25b), *roubaram e estragaram* é um predicado; em (26b), *simples e competente* é um pré-núcleo. Esse parece, pois, ser um processo de ampla atuação na língua, e reconhecê-lo nos permite evitar a repetição de muitas funções; reservamos a noção de "repetição" para algumas funções que podem ocorrer mais de uma vez **sem coordenação**. É bom notar que o princípio, embora muito geral, não é isento de exceções: a negação verbal e o predeterminante, por exemplo, não admitem coordenação.

Nesse caso, então, seguindo o princípio geral da coordenação, O_1 de (23) tem a mesma função que O_2 ou O_3, caso estas ocorram sozinhas. E como a função determina a classe, O_1 deve ser uma oração, tal como O_2 e O_3.

Uma nota: em (25), não seríamos obrigados a considerar, então, *roubaram e estragaram* como um verbo, já que tanto *roubaram* quanto *estragaram*, isolados, são verbos? Em tese, isso deveria ser feito, e há lingüistas que o fazem. Mas a separação estrita entre morfologia e sintaxe, adotada nesta análise, nos impede de colocar uma seqüência de palavras na mesma classe que uma palavra isolada. Por isso, e só por isso, não chamarei a seqüência *roubaram e estragaram* de "verbo", nem a seqüência *simples e competente* de "adjetivo". Mas as razões sintáticas para colocá-las nessas classes, sem dúvida, existem.

Resta um problema a resolver: se a estrutura de (3) é (23), qual é a relação entre a oração maior, O_1, e as duas orações menores, O_2 e O_3? Alguns autores, como Huddleston, consideram O_2 e O_3 subordinadas de O_1, e talvez essa seja a melhor solução. Mas não há dúvida de que se trata de um caso muito peculiar de subordinação: que funções desempenham O_2 e O_3 dentro de O_1?

Vou adotar a solução sugerida por Huddleston, definindo uma função especial para exprimir a relação entre elementos coordenados (não apenas orações) e o termo que os engloba: chamarei essa função **membro de coordenação** (MC). Assim, O_2, por exemplo, é MC de O_1, o que justifica chamá-la de "subordinada" de O_1. E igualmente o termo *simples* e o termo *competente* em (26b) são MCs do termo *simples e competente*, cuja função é a de pré-núcleo.

Essa análise tem uma conseqüência algo surpreendente: não existe coordenação sem subordinação — já que, por definição, se A e B são coordenados, eles (em geral com o acréscimo de uma conjunção) formam um terceiro constituinte C, ao qual A e B são subordinados, por terem em C a função de membro de coordenação.

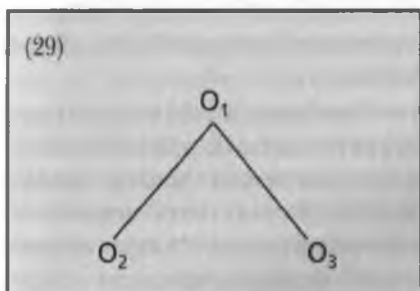
5.2.2.4. Coordenação de subordinadas?

Vimos na seção precedente que muitos termos de uma oração podem aparecer coordenados. Isso acontece mesmo quando os termos contêm orações, isto é, quando são oracionais. Assim, na frase

(27) Flit percebeu que não havia saída.

o objeto direto de *percebeu* é o sintagma oracional *que não havia saída*, que contém a oração *não havia saída*. É possível formar um sintagma oracional contendo duas orações coordenadas, como em

(28) Flit percebeu que não havia saída e o jogo estava perdido.

A análise tradicional vê em (28) duas subordinadas coordenadas entre si. Mas essa talvez não seja a melhor análise: relembro a hipótese levantada anteriormente, de que duas orações coordenadas se articulam em uma oração maior, da seguinte forma:

Nesse caso, o que temos em (28) é **uma** oração subordinada, que por sua vez contém duas coordenadas.

A subordinada é O_2: *não havia saída e o jogo estava perdido.* Nesse caso, não diremos que O_3 (*não havia saída*) e O_4 (*o jogo estava perdido*) são subordinadas de O_1, mas antes que são as duas coordenadas que compõem O_2 (sobre o estatuto exato de O_3 e O_4 dentro de O_2, ver 5.2.2.3. acima). Há portanto **quatro** orações em (28), sendo uma principal (O_1: *Flit percebeu que não havia saída e o jogo estava perdido*); uma subordinada a O_1 (O_2: *não havia saída e o jogo estava perdido*); e duas coordenadas (O_3: *não havia saída* e O_4: *o jogo estava perdido*).

5.2.3. Marcas de subordinação

Em uma oração complexa composta por subordinação, é necessário que haja uma marca qualquer para distinguir a oração subordinada.

A própria estrutura em geral seria suficiente para isso: em uma sentença como

(1) Titia disse que nós desarrumamos a casa.

só é possível distinguir uma oração com todos os seus termos preenchidos por sintagmas não-oracionais, a saber, *nós desarrumamos a casa.* A outra oração, *titia disse que nós desarrumamos a casa,* teria que compreender um objeto direto para o verbo *dizer* (que exige OD), e o único candidato possível é o sintagma *que nós desarrumamos a casa,* que é oracional.

Isso acontece em geral quando há orações subordinadas e pode ser considerado uma marca de subordinação: uma oração com todos os seus termos preenchidos por sintagmas não-oracionais não pode ser principal; e uma oração que tenha pelo menos um termo preenchido por sintagma oracional é uma oração principal. Esse critério, por si só, bastaria para distinguir subordinadas de principais em muitos casos.

No entanto, a língua não se contenta com isso; em todos os casos, as orações subordinadas são marcadas pela presença de certos elementos sintáticos ou morfológicos, que a caracterizam como subordinada, embora nem sempre sejam propriamente parte da subordinada. Esses elementos têm também uma função semântica (isto é, significam alguma coisa), de modo que não seriam dispensáveis, mesmo em princípio.

A lista das marcas explícitas de subordinação é relativamente curta:

(a) conjunções (subordinativas);

(b) relativos;

(c) marcas de interrogativa indireta;

(d) desinências do modo subjuntivo;

(e) desinências de gerúndio e de infinitivo.

Nos casos (a), (b) e (c), as marcas são formadas de palavras mais ou menos autônomas (marcas sintáticas); são "formas livres" segundo a terminologia usual ("formas dependentes" segundo a distinção de Camara, 1964). Nos casos (d) e (e), as marcas são morfemas presos (marcas morfológicas). A seguir, vou exemplificar brevemente o funcionamento dessas marcas na caracterização das orações subordinadas.

Antes, uma observação importante: a função sintática dessas marcas pode ser interpretada como a de formar sintagmas a partir de orações. Assim, *que* mais uma oração formam um sintagma nominal; e uma oração cujo verbo tem a desinência de gerúndio funciona como um sintagma "adverbial". Embora essa observação deixe de exprimir certos aspectos do fenômeno (por exemplo, uma oração de gerúndio não possui tempo morfologicamente marcado), é importante para que se possam manter generalizações como a de que "a função de sujeito é sempre desempenhada por um sintagma nominal".

5.2.3.1. *Conjunções*

As conjunções que nos interessam no momento são subordinativas, porque sua função é a de inserir uma oração (a subordinada) dentro de outra oração (a principal). Na presente análise, o termo **conjunção** se aplica apenas às "conjunções subordinativas" da gramática tradicional; as "conjunções coordenativas" serão denominadas **coordenadores**. Assim, podemos dizer que as conjunções são itens léxicos que, colocados imediatamente antes de uma oração, formam com ela um sintagma que é termo de alguma oração maior (no capítulo 13, o leitor encontrará um estudo mais detalhado das conjunções). Uma conjunção pode ser apresentada graficamente como uma palavra (*que, quando*) ou como um grupo de palavras (*visto que, se bem que, sempre que*). Conjunções deste último tipo funcionam de certo modo como palavras únicas, pois suas partes não têm independência sintática; e, a se levar em conta apenas a sintaxe, poderiam ser grafadas sem espaço, como uma única palavra. No entanto, a ortografia reflete o fato de que as partes dessas conjunções têm alguma independência fonológica, pois conservam cada uma seu acento tônico próprio.

A classificação tradicional das conjunções "subordinativas" compreende, primeiro, a oposição entre "conjunções adverbiais" e "conjunções integrantes"; e, depois, a subclassificação das conjunções adverbiais em "causais", "comparativas", "temporais" etc.

A primeira dessas oposições ("adverbiais" X "integrantes") tem base sintática: as conjunções adverbiais são aquelas que, seguidas de uma oração, formam um "sintagma adverbial" — ou seja, um sintagma que pode desempenhar funções como a de adjunto circunstancial, atributo ou adjunto oracional. No entanto, não vou adotar essa subclassificação, porque na análise adotada nesta *Gramática* essas funções não são englobadas sob o rótulo de "funções adverbiais" (a não ser para uma referência informal).

Já a subclassificação das conjunções adverbiais em "causais", "temporais" etc. tem base exclusivamente semântica e não caberia neste capítulo. E, de todo modo, a classificação tradicional é excessivamente pobre para exprimir toda a variedade de relações semânticas que as conjunções podem veicular. Por conseguinte, prefiro não subclassificar aqui as conjunções subordinativas, limitando-me a algumas sugestões como as que se seguem.

As conjunções, como vimos, colocam-se logo antes de uma oração, e o conjunto tem propriedades sintáticas semelhantes às de certos sintagmas não-oracionais. Vimos, por exemplo, que a seqüência *que nós desarrumamos a casa* (= *que* + oração) tem potencial funcional semelhante ao de um SN. E uma seqüên-

cia como *quando Ana Maria me chamou* funciona de maneira muito semelhante à do sintagma não-oracional *às duas horas em ponto*:

(31) Entrei no apartamento às duas horas em ponto.

(32) Entrei no apartamento quando Ana Maria me chamou.

Em ambas as frases, temos um adjunto circunstancial: na primeira frase, este é um sintagma não-oracional; na segunda, é uma seqüência de uma conjunção mais uma oração.Os sintagmas oracionais podem, assim, desempenhar diversas funções: objeto direto, adjunto circunstancial, sujeito, atributo etc. (ver a seção 5.3., onde dou a lista completa). E as conjunções, pelo menos em parte, são especializadas na introdução de sintagmas com determinada função. Assim, *quando* nunca introduz um objeto direto nem um sujeito.

Temos aí um critério potencialmente interessante para subclassificar as conjunções. É preciso notar, desde já, que algumas podem introduzir mais de um tipo de sintagma (*que*, talvez a mais versátil das conjunções, parece poder introduzir SNs e também alguns tipos de sintagmas "adverbiais").

5.2.3.2. Relativos

Relativos são elementos introdutores de uma construção especial, a **construção relativa**; essa construção também recebe a designação tradicional de "oração adjetiva". Um exemplo é

(33) O gato que lambeu meu queijo era angorá.

A construção relativa é *que lambeu meu queijo*. Essa seqüência forma uma oração subordinada, cuja função é a de modificador. Trata-se, portanto, de um constituinte de nível suboracional, componente do SN *o gato que lambeu meu queijo*.

Nesse caso, o caráter subordinado da oração é marcado pela estrutura peculiar à construção relativa; esta se caracteriza por:

(a) presença de um relativo (os relativos são *que*, *o qual*, *quem*, *onde*, *cujo*), precedido às vezes de uma preposição;

(b) presença de uma estrutura oracional **aparentemente incompleta**, logo após o relativo;

(c) articulação de um **elemento nominal** (parte de um SN) + **o relativo** + **a estrutura oracional mencionada**, formando uma seqüência que é um SN; o elemento nominal inicial nem sempre está presente.

Assim se caracteriza uma construção relativa; e esses traços são suficientemente específicos para tornar fácil a identificação da construção. Vamos examinar alguns exemplos.

Seja, inicialmente, o exemplo:

(33) O gato que lambeu meu queijo era angorá.

Como essa sentença contém dois verbos em forma finita, deve conter duas orações. Isso é confirmado pelas marcas características da construção relativa, a saber: (a) presença do relativo *que*, (b) presença da seqüência *lambeu meu queijo*, que vem logo após o relativo e tem a estrutura de uma oração incompleta (falta o sujeito); e (c) presença da seqüência *o gato que lambeu meu queijo*, que é formada de um elemento nominal (*o gato*), seguido de *que*, seguido da estrutura oracional, e que é um SN (no caso, é o sujeito de *era*).

Outro exemplo é a frase

(34) O estrago que o gato fez ficou sem conserto.

Aqui temos: (a) o relativo *que*; (b) a oração (aparentemente) incompleta *o gato fez*, à qual faltaria o objeto direto; e (c) a seqüência *o estrago + que + o gato fez*, que é um SN (sujeito de *ficou*).

Um terceiro exemplo pode ser

(35) O bicho do qual eu falei é um gato.

onde encontramos: (a) o relativo *o qual*, aqui precedido da preposição *de*; (b) a estrutura oracional *eu falei*; e (c) o SN *o bicho + do qual + eu falei*.

A estrutura nominal inicial do SN que contém a construção relativa nem sempre aparece, conforme foi adiantado acima. Nesses casos, a construção relativa se inicia diretamente com o relativo (com ou sem preposição), como em

(36) Ele fala mal de quem o ajuda.

Mas ainda aqui os traços característicos da construção relativa estão presentes: (a) o relativo *quem*, precedido da preposição *de*; (b) a estrutura oracional *o ajuda*, aparentemente sem sujeito; e (c) o SN formado de *quem + o ajuda*.

As construções relativas serão estudadas na seção 5.3.3.1. Aqui, basta-nos observar que o relativo, juntamente com os demais traços da construção relativa, funciona como marca de subordinação. A construção relativa, como se verá, constitui uma oração; como tem uma função dentro de outra (a de modificador), é sempre uma subordinada.

5.2.3.3. Marcas de interrogativa indireta

Interrogativas indiretas, ou mais corretamente **interrogativas subordinadas**, se parecem superficialmente com as relativas; mas há diferenças relevantes, que serão vistas na seção 5.3. Aqui trataremos apenas dos elementos que as marcam, que são também em parte idênticos aos que marcam as relativas. Um exemplo de interrogativa indireta é

(37) Não sei quantos espectadores virão ao circo.

Segundo os critérios estabelecidos aqui, essa sentença tem duas orações, sendo que a subordinada (*quantos espectadores virão ao circo*) funciona como objeto direto de *sei*. Essa subordinada (uma interrogativa indireta) é marcada pelo elemento introdutor *quantos*. Os itens que funcionam como marcas de interrogativas indiretas são: *o que, quem, onde, quando, quanto(s), como, qual*. Seguem-se alguns exemplos:

(38) Imagino quem faria uma coisa dessas.

(39) O chefe ignora quando teremos uma folga.

(40) Não imagino como você vai conseguir esse emprego.

Os elementos introdutores têm função dentro da subordinada, e portanto fazem parte dela. Assim, em (37), *quantos* é determinante dentro do SN *quantos espectadores*; em (38), *quem* é sujeito de *faria*; e, em (39), *quando* é atributo de *teremos*.

5.2.3.4. Desinência de subjuntivo

Em uma frase como

(41) A empresa deseja que todos tenham uma boa viagem.

a subordinada (*todos tenham uma boa viagem*) é introduzida por nossa conhecida conjunção *que*; mas, além disso, o NdP é um verbo no subjuntivo. O subjuntivo pode ser também considerado uma marca de subordinação, porque só em certas estruturas excepcionais, e pouco freqüentes (ver 3.1.3.), pode ocorrer em uma oração não-subordinada.

Além disso, em alguns poucos casos, o subjuntivo pode ser a única marca explícita de subordinação, como em

(42) O requerente solicita seja feita uma devassa.

Aqui, a ocorrência de *que* antes da subordinada é opcional.

5.2.3.5. Desinências de infinitivo e de gerúndio

Finalmente, temos as desinências de infinitivo e de gerúndio, que também marcam subordinadas. Em frases como

(43) Daniel trabalha vendendo cimento.

(44) Daniel resolveu vender cimento.

as subordinadas, respectivamente *vendendo cimento* e *vender cimento*, são marcadas como tais pela forma do verbo. O infinitivo e o gerúndio são formas especializadas na ocorrência em orações subordinadas, e portanto marcam eficientemente suas orações como subordinadas. Não há orações não-subordinadas com gerúndio, e as não-subordinadas com infinitivo são raras (ocorrem em fra-

ses como *olhar à direita!*, interpretadas como imperativas).

5.2.4. Marcas de coordenação

A subordinação é um fenômeno estrutural, em última análise bastante simples. É decorrência direta do modo como se estrutura a oração, acrescido à possibilidade de se encaixar orações dentro de diversos tipos de sintagmas. Uma oração subordinada, assim como o sintagma de que faz parte são constituintes de pleno direito da oração. Em certo nível de análise, uma oração complexa é em tudo idêntica a uma oração simples.

Já os casos tradicionalmente classificados como de coordenação não foram um grupo tão consistente e homogêneo. A união observada entre orações coordenadas é de natureza muito menos estreita do que a que existe entre uma principal e suas subordinadas. Em muitos casos, a coordenação se aproxima dos fenômenos discursivos, muito menos dependentes da estrutura interna das formas lingüísticas e baseados, em vez disso, de preferência em fatores semânticos e cognitivos em geral.

Assim, um tratamento unificado da coordenação dentro da sintaxe é praticamente impossível. Nesta seção, vou descrever as marcas formais que caracterizam as estruturas coordenadas, subdividindo-as em pelo menos três grupos, que apresentam pouca afinidade uns com os outros; às marcas, será preciso acrescentar a justaposição pura e simples das orações, em geral com algum sinal de pontuação entre elas. Quase a única afirmação que vale para todos os casos é a de que, quando duas ou mais estruturas são unidas por coordenação, nenhuma delas exerce função sintática dentro de nenhuma outra: definição negativa e que vale para outras relações que não a de coordenação (por exemplo, para duas funções diferentes, digamos o sujeito e o objeto direto, da mesma oração).

5.2.4.1. Coordenação sem marca

Ao contrário da subordinação, a coordenação pode se dar sem nenhuma marca explícita, simplesmente justapondo duas orações. Nesses casos, coloca-se sempre algum sinal de pontuação, como a vírgula:

(45) Seu bode comeu minha camisa, você vai pagar o prejuízo.

A marca de coordenação, propriamente dita, é apenas o sinal de pontuação. Na fala, há um contorno entonacional característico, com entonação descendente, às vezes acompanhada de pausa, no final da primeira oração. No caso da escrita, se a marca for uma vírgula, o único cri-

tério seguro de identificação das orações é a estrutura interna. Em (45), então, o que nos diz realmente que há duas orações é apenas a presença de duas estruturas oracionais completas, lado a lado.

A gramática tradicional chama "assindéticas" as orações coordenadas sem conjunção; e "sindéticas" as que têm conjunção.

5.2.4.2. Coordenação com *e*, *ou* e *mas*

Examinemos agora o funcionamento do coordenador ("conjunção coordenativa") *e*, que ilustra a coordenação em sua forma menos problemática. Um exemplo já visto é

(3) Titia fez a salada e mamãe fritou os pastéis.

Algumas propriedades da coordenação com o coordenador *e* são as seguintes:

(a) As duas orações podem ser separadas, opcionalmente, por pontuação:

(46) a. Titia fez a salada; e mamãe fritou os pastéis.
b. Titia fez a salada. E mamãe fritou os pastéis.

Aqui já se vê como é relativamente fraca a conexão entre os membros da coordenação: em (46), já nem sequer temos um período único, mas duas unidades que podem (e talvez devam) ser tratadas independentemente pela análise sintática.

(b) O coordenador vale para coordenar qualquer número de membros; nesse caso, em geral, mas não obrigatoriamente, o coordenador só ocorre entre os dois últimos membros:

(47) Marcelo adormeceu, Patrícia saiu e Maria foi ver TV.

(c) O coordenador, quando não repetido, só pode ocorrer em uma posição, ou seja, logo antes da última oração:

(48) * Titia fez a salada, mamãe e fritou os pastéis.

(d) Pode-se acrescentar que *e* pode servir para juntar quaisquer elementos coordenáveis; assim, encontramo-lo coordenando orações, como em (1), ou SNs, como em

(24) b. Avistei hoje seu patrão e o doutor Márcio.

ou verbos, como em

(25) b. Vocês roubaram e estragaram meu chocalho.

ou ainda adjetivos, como em

(26) b. Um simples e competente auxiliar

Essas são características da concepção tradicional de coordenação. No entanto, somente dois itens as apresentam todas, a saber, *e* e *ou*. Todos os outros coordenadores têm comportamento diferente. Toman-

do inicialmente *mas*, verificamos que não possui a propriedade (b): só pode coordenar dois elementos, nunca mais de dois:

(49) Carolina chegou mas não fez declarações.
(50) Carolina chegou, desembarcou, mas não fez declarações.

Pode-se verificar que em (50) o coordenador *mas* não vale para coordenar as três orações, porque é possível introduzir *e* entre as duas primeiras.

(51) Carolina chegou e desembarcou, mas não fez declarações.

A interpretação semântica confirma isso, pois não há relação adversativa entre *Carolina chegou* e *desembarcou.*

Além disso, *mas* não pode tampouco coordenar quaisquer elementos coordenáveis: podemos encontrar *mas* ligando adjetivos ou verbos, mas não SNs.

Em resumo, *e* e *ou* integram o primeiro grupo de coordenadores em que *mas* figura como um terceiro membro, bastante desviante.

5.2.4.3. Porém

Também são considerados casos de coordenação orações ligadas por uma série de outros itens, cujo comportamento sintático é muito diferente do de *e, ou*. São exemplos desses itens: *porém, entretanto, no entanto, contudo, não obstante, todavia, senão, assim, por isso, por conseguinte, de modo que, em vista disso* e vários outros. Esse conjunto certamente não é homogêneo; mas, como suas diferenças de comportamento não foram ainda devidamente estudadas, vou limitar-me aqui a uma descrição preliminar da sintaxe de *porém,* com alguns comentários sobre um ou outro membro mais aberrante do grupo.

Vamos comparar *porém* com *e,* baseando-nos nos quatro traços vistos na seção precedente.

(a) Orações coordenadas com *porém* também aceitam pontuação; mas, ao contrário do que acontece com *e,* a pontuação é obrigatória; e o ponto e vírgula parece ser a opção mais aceitável.

(52) a. ?? Titia adormeceu porém vovó continuou a cantar.
b. Titia adormeceu, porém vovó continuou a cantar.
c. Titia adormeceu; porém vovó continuou a cantar.

É algo difícil fazer julgamentos de aceitabilidade precisos com esses exemplos. Para mim, pelo menos, (52a) é nitidamente a pior das três e (52b) é ainda um pouco pior do que (52c).

(b) Tal como todas as conjunções, exceto *e* e *ou, porém* só pode coordenar dois elementos, nunca mais de dois.

(c) *Porém*, ao contrário de *e*, *ou* e *mas*, pode ocorrer em outras posições que não logo antes da segunda oração coordenada:

(53) a. Titia adormeceu; porém vovó continuou a cantar.
b. Titia adormeceu; vovó, porém, continuou a cantar.
c. Titia adormeceu; vovó continuou, porém, a cantar.
d. Titia adormeceu; vovó continuou a cantar, porém.

Essa propriedade distingue *porém* de todos os coordenadores vistos até o momento. Voltarei a ela logo adiante.

(d) *Porém*, tal como *mas*, não pode coordenar sintagmas nominais.

A mais surpreendente dessas propriedades é a de poder ocorrer em diversas posições na oração; com exceção deste grupo, todos os coordenadores têm posição fixa, ou no início da última oração coordenada, ou (no caso dos descontínuos, ver 5.2.4.4.) no início de cada um dos elementos coordenados. A possibilidade de movimentação na oração aproxima *porém* de certos itens usualmente analisados como "advérbios"; assim, podemos ter

(54) a. Titia adormeceu; conseqüentemente, vovó parou de cantar.
b. Titia adormeceu; vovó, conseqüentemente, parou de cantar.

e assim por diante: *conseqüentemente* pode ocorrer nas mesmas posições que *porém*.

Esse fato, juntamente com a observação de que coordenadas podem ocorrer sem coordenador, permite levantar a hipótese de que itens como *porém* não seriam em absoluto coordenadores, mas advérbios da classe de *conseqüentemente*. Observe-se que o significado de *conseqüentemente* é tal que seu uso estabelece uma conexão semântica entre as duas orações; poderia ser a mesma coisa com *porém*. Nesse caso, frases como as de (53), acima, seriam exemplos de coordenação sem coordenador (assindética).

Note-se que o grupo de advérbios cujo comportamento sintático se assemelha ao de *porém* é grande e inclui formas complexas como, por exemplo, *apesar disso*, *de tal maneira que* etc. Em geral, não se encontra, em classes de "palavras gramaticais", como as conjunções, coordenadores, preposições, artigos, pronomes, tal profusão de elementos. Não seria uma indicação do caráter adverbial de *porém* — que, nesse caso, pertenceria a uma classe aberta?

Isso deve ficar, por ora, como simples hipótese; somente a pesquisa poderá esclarecer a questão. Como sugestões para o direcionamento dessa pesquisa, podemos formular perguntas como:

Há diferenças significativas entre *porém* e *conseqüentemente*, que

permitam separá-los em classes diferentes?

As propriedades de *porém* o fazem parecer mais com os coordenadores do que com os advérbios? E, se a resposta é positiva, quais dos outros itens geralmente classificados como "conjunções coordenativas" (aqui, **coordenadores**) pertencem ao grupo de *porém*?

Outra pergunta de interesse teria a ver com as possibilidades de coocorrência de *porém* com advérbios de determinada classe. Por exemplo, *porém* não coocorre com *conseqüentemente*:

(55) ?? Titia dormiu; porém conseqüentemente vovó parou de cantar.

Essa frase é pelo menos estranha. Pergunta-se: as condições que fazem (55) inaceitável (ou pouco aceitável) têm a ver com as condições de coocorrência de advérbios de certa classe entre eles?

Finalmente, é importante observar que esses "coordenadores", ao contrário de quaisquer outros, podem coocorrer com marcas de subordinação dentro da mesma oração. Assim, é possível dizer

(56) Consegui um empréstimo, que porém não me ajudou a sair do déficit.

Aqui, *porém* coocorre com um relativo. Encontra-se ainda *porém* (e outros membros deste grupo) mesmo ao lado de coordenadores, como em

(57) Trabalhamos duro; e porém não fomos reconhecidos.

Esse comportamento é difícil de compatibilizar com uma análise que faz de *porém* um coordenador.

O leitor terá observado que me referi várias vezes a "advérbios de certa classe"; é que o problema está entrelaçado com o da classificação das palavras tradicionalmente chamadas "advérbios". Essa é uma das classes mais mal definidas da gramática tradicional; e é certo que o rótulo "advérbios" engloba mais de uma classe. A questão da classificação de *porém* não pode ser separada da questão mais ampla da classificação dos "advérbios", sendo mesmo parte dela. Aqui não é possível levar mais adiante a discussão, que deverá aguardar resultados de novas pesquisas (ver mais algumas observações sobre os "advérbios" na seção 13.5.).

Nesta seção, acabei de examinar os traços sintáticos de *porém*, considerando essa palavra como representante de todo um grupo de assim chamados "coordenadores". Mas relembro que o grupo é heterogêneo; qualquer generalização depende do estudo individual do comportamento sintático de cada um de seus integrantes. Outros membros desse grupo são: *entretanto, no entanto, senão, todavia, não obstante, então, assim, por isso, em vista disso, de maneira que, por conseguinte, contudo* etc.

5.2.4.4. *Conectivos descontínuos*

Há um grupo de elementos que funcionam como conectivos de orações (assim como de outros tipos de constituintes) e que são compostos de dois elementos, que sempre coocorrem, mas não contiguamente. O exemplo prototípico é *não só... mas também* (e suas variantes: *não só... mas ainda*; *não apenas... mas também*; *não só... como ainda* etc.):

(58) Joaquim não só toca viola mas também canta fados.

Esses itens possuem a maior parte das características que marcam *e* e *ou* como "verdadeiros" coordenadores: a pontuação é opcional (em (58), uma vírgula pode ocorrer antes de *mas*); a posição do coordenador é fixa (a primeira parte precede o primeiro elemento coordenado, e a segunda precede o segundo); e funcionam para quaisquer constituintes coordenáveis, inclusive sintagmas nominais:

(59) Joaquim vendeu não só o armazém mas também a loja.

(60) Não só Joaquim, mas também Manuel abriu falência.

Os descontínuos só diferem de *e* por não poderem coordenar mais de dois elementos. Portanto, as indicações são de que seriam conectivos descontínuos, diferindo de *e* basicamente por marcarem **ambos** os membros da coordenação. Outros exemplos são: *tanto... quanto*; *ora... ora*; *já... já*; *quer... quer.*

Como se vê, alguns desses conectivos descontínuos são formados pela repetição do mesmo elemento. E alguns deles podem também ocorrer sem repetição; assim, temos *ou... ou* e *nem... nem*, ao lado de *ou* e *nem* sem repetição. Talvez a melhor análise nesses casos seja considerar que certos coordenadores podem, opcionalmente, ocorrer repetidos, diante de ambos os membros da coordenação (caso de *ou*), e que alguns são excepcionais por só ocorrerem repetidos (caso de *quer... quer*). Assim, os únicos conectivos realmente descontínuos, isto é, não formados pela repetição de um elemento diante dos constituintes coordenados, seriam *tanto... quanto, tanto... como* e as várias formas de *não só... mas também.*

5.3. FUNÇÕES E ESTRUTURA DOS SINTAGMAS COMPLEXOS

Chamaremos **sintagma complexo** um sintagma que encerra, como um de seus constituintes imediatos ou mediatos, uma oração. Assim, o SN

(61) que nós desarrumamos a casa

que, como vimos, é o objeto direto de

(1) Titia disse que nós desarrumamos a casa.

é complexo, porque um de seus constituintes é a oração *nós desarrumamos a casa.* Em princípio, um sintagma complexo pode ocupar as mesmas funções que um sintagma simples da mesma classe (pois, de outra forma, não se justificaria colocá-los na mesma classe). Assim, um SN simples pode ser sujeito, objeto direto etc., e um SN complexo pode desempenhar exatamente as mesmas funções.

Por outro lado, algumas funções só podem ser desempenhadas por sintagmas simples; por isso, é necessário dar uma lista das funções que podem ser desempenhadas por sintagmas complexos. Além disso, a estrutura interna dos sintagmas complexos difere bastante, segundo a classe a que pertence: um SN complexo não se estrutura da mesma maneira que um SAdj complexo. É o que passaremos a ver.

5.3.1. Sujeito e objeto direto

Entre as funções que podem ser preenchidas por sintagmas complexos estão as de **sujeito**, **objeto direto**, **adjunto circunstancial**, **complemento do predicado**, **adjunto oracional** e **atributo** (no nível da oração) e a de **modificador externo** (no nível do sintagma). Outras funções, como o **predicativo**, a **negação verbal**, o **núcleo do predicado**, assim como o **determinante**, o **quantificador**, os **pré-núcleos** etc., não podem nunca ser ocupadas por sintagmas complexos. Darei exemplos dos casos mais freqüentes, descrevendo sua estrutura. Mas é possível que haja mais algumas funções que podem ser ocupadas por sintagmas complexos; falta um levantamento completo.

Em (1), como sabemos, temos um objeto direto complexo:

(1) Titia disse que nós desarrumamos a casa.

O mesmo sintagma poderia ocorrer, como qualquer SN, na função de sujeito:

(62) Que nós desarrumamos a casa é evidente.

A gramática tradicional chama a seqüência *que nós desarrumamos a casa* de "oração objetiva direta" [no caso de (1)] e "oração subjetiva" [no caso de (62)]. Prefiro não adotar essa nomenclatura porque, como vimos na seção 5.2.2.2., *que nós desarrumamos a casa* é um SN, e não uma oração; a oração é apenas *nós desarrumamos a casa* — mas não é esse o objeto direto de (1) nem o sujeito de (62); o objeto direto de (1) e o sujeito de (62) incluem a palavra *que.* Assim, se for necessário distinguir esse constituinte de um SN simples na mesma função, será melhor chamá-lo **objeto direto complexo**.

As funções de sujeito e objeto direto são, como sabemos, desempenhadas por sintagmas nominais; portanto, no caso de sujeito ou objeto direto complexos, por SNs complexos. Um SN complexo pode ter quatro tipos de estrutura.

Primeiro, pode ser composto de uma conjunção, em especial *que* ou *se*, mais uma oração com verbo no indicativo ou no subjuntivo; por exemplo,

(1) Titia disse que nós desarrumamos a casa.

(63) Titia ignora se nós desarrumamos a casa.

Em segundo lugar, pode ser composto de uma oração com verbo no infinitivo; por exemplo,

(64) Titia resolveu arrumar a casa.

(65) Titia lamenta Marcinha estar namorando.

O SN complexo pode ainda ser uma construção relativa sem antecedente:

(66) Quem colher mais café ganhará uma xícara.

(As construções relativas serão estudadas na seção 5.3.3.1.)

Finalmente, o SN complexo pode ser uma interrogativa indireta (ver 5.2.3.3.):

(67) Titia não sabe quem desarrumou a casa.

(68) Estou investigando quantos gatos cabem em um fusca.

Esses são os quatro recursos que a língua oferece para inserir uma oração dentro de um SN. Naturalmente, o uso de um ou outro depende de diversos fatores, raramente sendo livre. Assim, o verbo que ocupa o NdP pode determinar se se pode ou não usar um OD no infinitivo, ou, usando conjunção, se o modo do verbo da subordinada será subjuntivo ou indicativo (ver a seção 6.4.). O uso de relativas sem antecedentes, frente a interrogativas indiretas ou a orações com *que* ou infinitivo, é ditado pela semântica de cada uma; isso se pode ver pela diferença de significado entre as frases

(67) Titia não sabe quem desarrumou a casa.

(69) Titia não sabe que eu desarrumei a casa.

5.3.2. Atr, AA, AO e AC

Essas funções são desempenhadas por "sintagmas adverbiais" (ver 4.6.). E esses sintagmas podem ser complexos, como no exemplo

(70) Zelinha sofreu muito quando Zelão se casou.

Neste caso, temos um atributo; os traços que definem essa fun-

ção estão bem claros nessa frase. Assim, o constituinte *quando Zelão se casou* é [+Ant]:

(71) Quando Zelão se casou, Zelinha sofreu muito.

É [+Cl]:

(72) Foi quando Zelão se casou que Zelinha sofreu muito.

É [+PA]:

(73) Zelinha, quando Zelão se casou, sofreu muito.

E é [–Q], como o leitor poderá facilmente verificar. Logo, trata-se de um atributo.

Um exemplo de adjunto adverbial complexo é

(74) A moça riu até ficar sem fôlego.

Não é difícil verificar que o constituinte *até ficar sem fôlego* tem os traços de um AA: [–CV, –Ant, –Q, +Cl, –PA].

Já em

(75) Conforme foi noticiado, os preços do abacaxi baixaram.

o sintagma complexo *conforme foi noticiado* é um adjunto oracional.

E temos um adjunto circunstancial complexo em

(76) André estuda ouvindo música.

É interessante notar que, das funções de nível oracional que podem incluir um SN, o predicativo é a única que não pode ser desempenhada por sintagma complexo; resta averiguar se isso é conseqüência de alguma restrição de ordem semântica.

5.3.3. Modificador externo

5.3.3.1. A construção relativa

Vimos no capítulo 4 que o SN compreende dois termos que se ordenam depois do núcleo: o modificador interno e o modificador externo. O segundo desses termos pode também ser representado por um sintagma complexo — no caso, naturalmente, um sintagma adjetivo. Um exemplo é

(77) O urso que me mordeu era branco.

O sujeito dessa frase é o SN *o urso que me mordeu*; esse SN inclui uma oração subordinada, *que me mordeu* — onde *que* é sujeito. Ou seja, o sujeito de (77) é um tipo de SN a que demos o nome de **construção relativa** (ver 5.2.3.2.). As orações subordinadas introduzidas por relativos têm sempre a função de modificador externo e pertencem à classe dos sintagmas adjetivos.

A estrutura interna da construção relativa é bastante peculiar e complexa. Seu estudo apresenta dificuldades, nem todas resolvidas; mas é possível dar uma boa idéia geral de como se organiza essa construção.

Como se viu em 5.2.3.2., a construção relativa tem as seguintes características:

(a) presença de um relativo, precedido às vezes de uma preposição;

(b) presença de uma estrutura oracional aparentemente incompleta, logo após o relativo;

(c) articulação de um **elemento nominal** (parte de um SN) + **o relativo** + **a estrutura oracional mencionada**, formando uma seqüência que é um SN; o elemento nominal inicial nem sempre está presente.

Em (77), o relativo é *que*; a estrutura oracional é *me mordeu*, que aparentemente não tem sujeito; e o conjunto *o urso que me mordeu* forma um SN.

A oração é só aparentemente incompleta, porque há razões para se analisar *que* como o sujeito de *mordeu*; assim, a oração subordinada é *que me mordeu*: ao contrário do que acontece com SNs complexos, nos SAdjs complexos, o conectivo (o relativo) faz parte da oração subordinada.

A razão principal para se analisar *que* como sujeito de (77) — seguindo, aliás, a análise tradicional — é a seguinte: embora *me mordeu* seja aparentemente uma oração sem sujeito, não se pode acrescentar um sujeito a ela:

(78) * O urso que ele me mordeu era branco.

O mesmo se repete para as demais funções que um relativo pode ocupar. Outro exemplo é

(79) O filme que Bebeto fez ganhou a Palma de Ouro.

Aqui temos, aparentemente, o verbo *fazer* sem objeto direto; ora, sabe-se que esse verbo nunca aparece sem OD (ver o capítulo 6 a respeito das exigências de complementos). Mas *fazer* em (79) não admite o acréscimo de um OD:

(80) * O filme que Bebeto o fez ganhou a Palma de Ouro.

Se admitirmos (com a gramática tradicional) que o relativo é objeto direto da subordinada, esses problemas desaparecerão: a subordinada é *que Bebeto fez*; o verbo *fazer* aparece aí como OD, o que está de acordo com sua transitividade; e não se pode acrescentar um objeto porque só pode haver um por oração. Assim, é bastante seguro concluir que o relativo ocupa uma função dentro da oração subordinada. O relativo propriamente dito funciona sempre como um SN (daí poder ser sujeito ou objeto); mas pode vir precedido de preposição e, nesse caso (como qualquer SN preposicionado), pode ocupar funções típicas de "sintagmas adverbiais"; é um AC em

(81) A garota a quem eu dei meu relógio não quer saber de mim.

Essas considerações se referem, claro, às funções do **relativo**; a oração relativa funciona sempre como modificador externo, sendo sempre parte de um SN.

Nos exemplos acima, o relativo desempenha funções de nível oracional: sujeito, objeto direto, adjunto circunstancial. Como se vê, o relativo (com a preposição que o precede, se houver) se posiciona obrigatoriamente no início de sua oração, qualquer que seja sua função. Semanticamente, ele se interpreta como correferente do elemento nominal que o precede (seu antecedente); assim, em (81), entende-se que eu dei o relógio à mesma garota que me despreza.

Existem, tal como previsto na definição de "construção relativa", casos em que não há antecedente precedendo o relativo:

(82) Só respeito quem o governo financia.

Fora isso, entretanto, *quem o governo financia* tem todos os traços de uma oração relativa: há um relativo, *quem* (que também pode ocorrer, com preposição, em orações relativas com antecedente); esse relativo desempenha uma função na subordinada (no caso, objeto direto); e o relativo se coloca no início da oração. Por isso, consideramos também *quem o governo financia* uma construção relativa — conseqüentemente, um SN, que no caso é coextensivo com a oração.

5.3.3.2. O relativo como modificador

Em todos os exemplos que vimos acima o relativo desempenha uma função de nível oracional: sujeito, objeto direto, adjunto circunstancial etc. No entanto, em um caso particular, encontra-se o relativo desempenhando uma função suboracional, a de modificador externo. Ou seja: dentro da construção relativa, que tem a função de modificador externo, o próprio relativo tem, por sua vez, também a função de modificador externo.

Esses casos têm estrutura parcialmente diferente da dos demais e demandam um estudo à parte. Trata-se, antes de tudo, das construções em que o relativo é *cujo*, como por exemplo

(83) O urso cuja pata eu cortei era branco.

Aqui a construção relativa, *cuja pata eu cortei*, é modificador externo de *urso*; e *cuja* é modificador externo de *pata*.

A razão para se analisar *cujo* como modificador externo é análoga à que se deu para as demais funções: não é possível acrescentar um modificador ao elemento *pata*:

(84) * O urso cuja pata dele eu cortei era branco.

O que este caso tem de especial é que:

(a) o relativo desempenha uma função dentro de um SN (isto é, uma função não-oracional);

(b) o relativo é colocado no início do SN a que pertence — ocupando, aparentemente, o lugar do determinante; e

(c) todo o SN em questão é colocado no início da oração.

O processo, como se vê, é mais complicado do que o que afeta os demais relativos.

Acredita-se que o relativo ocupa o lugar do determinante porque *cujo* nunca coocorre com determinante, embora possa coocorrer com outros termos do SN:

(85) a. * O urso cuja a pata eu cortei

b. * O urso cuja aquela pata eu cortei

c. O urso cujos poucos piolhos eu catei

d. O urso cujos quatro pés eu cortei

Em (85c), temos *cujo* com quantificador, e em (d) com numerador. Mas (85a) e (b), onde teríamos *cujo* com determinante, são mal formados.

Ainda a respeito de relativos em função de modificador, é preciso notar que há uma alternativa estrutural à construção com *cujo*. Nessa alternativa, usa-se um relativo comum (*quem* ou *o qual*), precedido de *de*; mas este não se coloca no início do SN, ficando no seu lugar normal após o núcleo do SN. E o SN inteiro se posiciona no início da oração, tal como ocorre na construção com *cujo*:

(86) O urso a pata do qual eu cortei era branco.

Essa construção me parece ser bem mais rara do que a alternativa com *cujo*.

Como vimos acima, uma das características da construção com relativo modificador é a colocação do SN inteiro no início da oração; assim, temos (83), mas não (87):

(83) O urso cuja pata eu cortei era branco.

(87) * O urso eu cortei cuja pata era branco.

Isso foi expresso acima dizendo-se que o SN a que o relativo pertence é colocado no início. Em certos casos, esse processo tem conseqüências curiosas. Quando o relativo pertence a um SN que por sua vez é parte de um segundo SN, pode-se transportar **qualquer um** dos SNs que contêm o relativo para o início da oração, resultando em grupos de sintagmas como

(88) a. O uniforme, do qual o Ministério especifica o feitio

b. O uniforme, o feitio do qual o Ministério especifica

O que está acontecendo aqui é que o SN a ser anteposto pode ser *qual* (com sua preposição, que também será anteposta), ou então *o feitio do qual*, que é também um SN e também contém o relativo. A língua per-

mite antepor qualquer um dos dois. O mesmo pode acontecer quando o relativo é *cujo*:

(89) a. O uniforme, de cujos bolsos o Ministério especifica o feitio
b. O uniforme, o feitio de cujos bolsos o Ministério especifica

Podemos mesmo ter três alternativas, sem cair em sintagmas tão complicados que se tornam improcessáveis:

(90) a. O uniforme, do qual o Ministério especifica o feitio dos bolsos
b. O uniforme, dos bolsos do qual o Ministério especifica o feitio
c. O uniforme, o feitio dos bolsos do qual o Ministério especifica

No exemplo (90), temos três SNs, um dentro do outro: *o feitio dos bolsos do qual*, que contém *os bolsos do qual*, que contém *o qual*. Em (89), só há dois SNs um dentro do outro, porque *cujos* é uma espécie de forma amalgamada de preposição + SN, que não pode ser considerada um SN. Além do mais, antepor *cujos* sozinho violaria a exigência (que não tem exceção) de que *cujo* se posicione sempre no início do SN que o contém:

(91) * O uniforme, cujo o Ministério especifica o feitio

É por isso que (89), essencialmente idêntico a (90), só admite duas versões, ao passo que (90) admite três.

Uma nota final, que pode ser de interesse: sabemos que, no português brasileiro coloquial, a construção relativa tem uma estrutura muito diferente da que foi exposta acima e é válida apenas para o padrão. Em particular, a construção com o relativo *cujo* praticamente desapareceu da língua falada, sendo substituída por uma construção regular do tipo

(92) O urso que eu cortei a pata dele era branco.

Correspondentemente, a construção com *cujo* nem sempre é usada com desenvoltura, mesmo quando a intenção é utilizar o português padrão; e as intuições de falantes instruídos a respeito dessa construção nem sempre são seguras.

5.3.3.3. Dois tipos de construção relativa

Tradicionalmente se distinguem as relativas em dois tipos: "explicativas" e "restritivas", nomes inspirados em suas propriedades semânticas. A diferença formal mais evidente entre elas é que as primeiras são sempre separadas por vírgula:

(93) O urso, que fugiu, era branco.
(94) O urso que fugiu era branco.

Na análise tradicional, (93) teria uma relativa explicativa, e (94) uma restritiva. Embora a análi-

se semântica subjacente a essas designações seja correta, prefiro não utilizar esses termos para não confundir o aspecto sintático com o semântico. Assim, chamaremos a relativa de (93) **apositiva**, e a de (94) **não-apositiva**.

As construções relativas apositivas têm estrutura sintática semelhante à das não-apositivas, com algumas diferenças, a saber:

(a) só as apositivas se separam por vírgula do resto da frase;

(b) só as apositivas podem ocorrer com o relativo *o qual* sem preposição; e

(c) só as apositivas admitem as construções múltiplas do tipo de (90), resultantes da movimentação de um SN que contém relativo modificador para o início da oração.

5.3.3.4. Uso dos diferentes relativos

Quanto ao uso dos diferentes relativos, é preciso observar o seguinte:

(a) *quem* só pode ocorrer sem antecedente [como em (82)], ou então precedido de preposição [como em (81)];

(b) *o qual* em construções relativas não-apositivas só ocorre precedido de preposição; seu uso sem preposição é restrito a construções relativas apositivas;

(c) *o que* só ocorre sem antecedente;

(d) finalmente, *que* só ocorre com antecedente; e pode ocorrer em construções apositivas ou não. *Que* ocorre também com preposições, mas sofre restrições algo misteriosas: com certas preposições pode ocorrer livremente, mas com outras é de aceitabilidade baixa, ou mesmo francamente inaceitável:

(95) a. A firma de que lhe falei

b. A firma a que me dirigi

c. ? A firma com que me entendi

d. ?? A firma a respeito de que lhe falei

e. * A firma contra que me revoltei

Essas restrições estão ainda por estudar.

5.3.4. Interrogativas indiretas

A interrogativa indireta (que talvez devesse, como já fiz notar, ser chamada **interrogativa subordinada**) constitui um tipo especial de sintagma complexo, com alguma semelhança superficial com a construção relativa, mas também com muitas diferenças relevantes. A estrutura da interrogativa indireta é essencialmente idêntica à de uma interrogativa não-subordinada (vista na seção 3.1.3.2.); essas orações podem subordinar-se sem modificar sua forma:

(96) Que roupa você vai usar?

(97) Não sei que roupa você vai usar.

Em (96), temos uma interrogativa não-subordinada (ou direta); em (97), a mesma estrutura aparece como subordinada.

À primeira vista, a estrutura parece semelhante à das relativas, principalmente por causa da presença de um elemento "Q" e do transporte do sintagma que o contém para o início da oração. No entanto, as diferenças são mais impressionantes: primeiro, os elementos interrogativos não são exatamente os mesmos que ocorrem como relativos; o sintagma interrogado não se reduz necessariamente ao elemento "Q"; e, principalmente, o sintagma complexo é um SN ou um "sintagma adverbial" (ao contrário da construção relativa, que como vimos é um SAdj). Vamos examinar cada um desses pontos.

Os elementos que introduzem a interrogativa indireta (e a interrogativa direta) são *que, qual, como, o que. Que* se coloca no lugar do determinante de um SN, que, portanto, não pode conter um dos determinantes usuais:

(97) Não sei que roupa você vai usar.

(98) * Não sei que a roupa você vai usar.

O comportamento de *qual* a esse respeito é curioso; ele parece coocorrer com o determinante, como em

(99) Não sei qual a roupa que você vai usar.

Mas as coisas não são simples: não é possível colocar *qual* + determinante simplesmente à testa de uma interrogativa indireta:

(100) * Não sei qual a roupa você vai usar.

Isto é, exige-se um segundo *que*, como está em (99).

Parece, então, que *qual* está interrogando apenas um SN, que inclui uma construção relativa. Não se trataria, portanto, de um caso de interrogativa indireta. Além disso, não é qualquer determinante que pode coocorrer com *qual*:

(101) * Não sei qual aquela roupa que você vai usar.

Por outro lado, *qual* pode ser usado sem determinante, e nesse caso se comporta de maneira idêntica a *que*:

(102) Não sei qual roupa você vai usar.

O interrogativo *como* pode também introduzir uma interrogativa indireta (mas nunca uma relativa):

(103) Não imagino como você vai sair dessa arapuca.

Os interrogativos *como* e *o que* são utilizados como sintagmas completos; já *que* e *qual* são determinantes de sintagmas maiores:

(97) Não sei que roupa você vai usar.

(104) Não sei o que você vai usar.

(105) * Não sei que você vai usar.

(106) * Não sei o que terno você vai usar.

Cujo nunca ocorre como introdutor de interrogativas indiretas.

Passando à segunda característica, acabamos de ver nos exemplos acima que o elemento interrogativo pode ser parte de um SN, o qual tem uma função dentro da oração subordinada. O interrogativo nunca requer um antecedente (como é o caso, em geral, dos relativos).

Finalmente, a interrogativa indireta é um sintagma nominal ou adverbial, ao contrário da construção relativa, que é um sintagma adjetivo. Assim, a interrogativa indireta é objeto direto (SN) em (97) e atributo ("sintagma adverbial") em (103). Pode ser também sujeito, como em

(107) Que roupa eu vou usar ainda é um problema.

O uso de interrogativas indiretas em função de sujeito parece ser um tanto raro.

6 Transitividade, regência e concordância

6.1. O FENÔMENO DA REGÊNCIA

As estruturas sintáticas se compõem de constituintes organizados em orações segundo certos princípios. Um desses princípios é o de que cada constituinte tem uma função sintática dentro da oração ou do sintagma a que pertence. As funções foram estudadas na seção 3.2.; vimos ali que cada função é definida por um tipo de comportamento sintático particular — assim, um sujeito se comporta sintaticamente de maneira diferente de um adjunto adverbial, por exemplo.

Agora vamos abordar outro aspecto da organização sintática das orações, que diz respeito, **grosso modo**, à propriedade de muitos itens léxicos de estipular certos traços da estrutura em que ocorrem. Assim, por exemplo, se construirmos uma oração cujo NdP é desempenhado pelo verbo *gostar*, teremos de incluir também um adjunto circunstancial precedido da preposição *de*. A ausência de tal preposição com o complemento causa inaceitabilidade:

(1) a. Todo gato gosta de sardinha.
b. * Todo gato gosta sardinha.

Esse fenômeno é tradicionalmente expresso dizendo-se que o verbo *gostar* **exige** a presença de *de* antes de seu complemento.

Essa relação é assimétrica, pois evidentemente não se poderia dizer que a preposição é que exige o verbo, já que *de* ocorre em muitas construções sem *gostar*. Diremos, então, que o verbo *gostar*, na frase (1a), **rege** o complemento (no sentido de que faz exigências quanto a sua presença e/ou sua forma); diz-se também que *gostar* é o **termo regente**, e o adjunto circunstancial com *de* é o **termo regido**. Temos aí, portanto, um exemplo do fenômeno da **regência**.

A relação termo regido/termo regente se manifesta de formas bastante variadas; em todos os casos, entretanto, pode-se distinguir um componente comum, a saber, um item léxico (termo regente) estipulando

certos traços da estrutura. A seguir, vou dar uma série de exemplos, que deverão deixar mais clara essa noção.

(a) Os verbos fazem exigências quanto à presença de certos termos em sua oração. Assim, o verbo *fazer* "exige" a presença de um objeto direto; tanto é assim que se pode dizer

(2) Gato faz barulho de noite.

mas não

(3) * Gato faz de noite.

Há também verbos que "recusam" certos termos: o verbo *nascer* não pode nunca ter objeto direto, ao passo que *morrer* pode, pelo menos em certos casos:

(4) * Ele nasceu um nascimento tranqüilo.
(5) Ele morreu uma morte tranqüila.

E há, naturalmente, muitos verbos que aceitam livremente a presença ou ausência de certos termos. Por exemplo, *comer* pode ocorrer com ou sem objeto direto, à vontade:

(6) Meu gato já comeu todo o mingau.
(7) Meu gato já comeu.

Este tipo particular de regência se denomina **transitividade verbal**, e será estudado na seção 6.2.

(b) Assim como os verbos, também itens de outras classes podem fazer exigências quanto à presença de complementos. A palavra *favorável* aceita um complemento do sintagma adjetivo (CSA), como em

(8) Minha decisão foi favorável à publicação do seu artigo.

Já o item *livresco*, por exemplo, não aceita complemento semelhante.

Outro exemplo é a palavra *favoravelmente*, que admite um complemento em

(9) Decidi favoravelmente à publicação do seu artigo.

ao passo que *estupidamente* nunca ocorre com complemento.

Em casos como os de *favorável*, *favoravelmente*, *livresco* e *estupidamente* falamos de **transitividade nominal**. Trataremos da transitividade nominal na seção 6.3.

(c) Tanto no caso da transitividade verbal quanto no da nominal, a exigência do termo regente não se limita sempre à presença de algum complemento. O termo regente pode estipular certos traços da forma desse complemento; tipicamente, o termo regente exige a presença de uma preposição específica. Assim, para manter os exemplos acima, o complemento de *gostar* vem com a preposição *de*; o de *comer* vem sem preposição; o de *favorável* vem com a preposição *a*, e assim por diante.

(d) Os verbos podem também fazer certas exigências concernentes à forma das orações que lhes são subordinadas. Essas exigências podem dizer respeito ao modo do verbo da subordinada, ou então ao tipo de subordinada: com *que* ou com infinitivo, por exemplo. Assim, uma oração que fa-

ça parte do objeto direto de *determinar* deve ser introduzida por *que* e ter o verbo no subjuntivo; não é aceitável um objeto direto com oração no indicativo ou infinitivo:

(10) a. Ele determinou que todos se calassem.
b. * Ele determinou que todos se calam/calaram.
c. * Ele determinou todos se calarem.

Já com o verbo *comentar* o objeto direto pode ocorrer com oração introduzida por *que* e indicativo, ou, então, com oração de infinitivo, mas não com oração de subjuntivo:

(11) a. O jornal comentou que o governo está perdido.
b. * O jornal comentou que o governo esteja perdido.
c. O jornal comentou o governo estar perdido.

Esse aspecto da regência dos verbos sobre suas subordinadas será visto na seção 6.4.

(e) Alguns itens não-verbais (tradicionalmente classificados como "advérbios") fazem exigência quanto ao modo da oração a que pertencem. Em geral, essa exigência é condicionada à posição do advérbio, de tal modo que o efeito só se faz sentir se o verbo vier **depois** do advérbio na oração:

(12) a. Eu talvez vá à Argentina este ano.
b. Eu irei à Argentina este ano, talvez.

Essas cinco categorias cobrem a maior parte dos fenômenos comumente denominados "de regência". Há outros casos, entretanto, que apresentam semelhanças com esses e poderiam, a rigor, ser considerados também como manifestações da relação de regência. Pode-se mencionar a rede de compatibilidades que funciona dentro (e em torno) do sintagma nominal. Por exemplo, a presença de um predeterminante anteposto exige a presença de um determinante, como em

(13) a. Todos os homens/todos esses homens
b. * Todos homens/* todos quatro homens.

Fenômenos de regência são muito numerosos na língua e nem sempre se reconhecem como tais. Nesta seção, vamos estudar de preferência aspectos da regência verbal (que é mais conhecida), com algum comentário sobre a regência nominal. Um estudo unificado dos fenômenos de regência está ainda por ser realizado.

6.2. TRANSITIVIDADE VERBAL

6.2.1. Crítica da classificação tradicional

Tradicionalmente, os verbos se distinguem em cinco tipos, de acordo com sua transitividade, a saber:

verbos transitivos diretos, transitivos indiretos, transitivos diretos e indiretos, intransitivos e de ligação. Conforme veremos, a classificação tradicional repousa sobre um equívoco fundamental e, conseqüentemente, não pode ser mantida. Será necessário propor um novo sistema.

A noção tradicional de verbo "transitivo" em oposição a "intransitivo" se define assim: um verbo é "transitivo" quando exige a presença de um objeto direto em sua oração; e é "intransitivo" quando recusa a presença de objeto direto. A definição é suficientemente clara, e dela decorre que sempre que houver em uma oração um verbo transitivo, essa oração deve ter objeto direto; e sempre que houver um verbo intransitivo, a oração não pode ter objeto direto. Note-se que o sistema não prevê lugar para verbos que possam ter OD ou não, à vontade; logo, é de se presumir que tais verbos não existam.

Na prática, porém, a definição não é respeitada. Classifica-se o verbo *comer* como "transitivo", porque aparece com OD em

(6) Meu gato já comeu todo o mingau.

Mas *comer*, como vimos, aparece igualmente sem OD:

(7) Meu gato já comeu.
(14) Meu gato quase não come.

Exemplos como esses — que, como veremos, são numerosos — colocam em xeque o sistema tradicional. A se seguir a definição dada, não haveria lugar para o verbo *comer*, que pode ocorrer com OD ou sem ele.

Procurando fugir ao problema, alguns autores sugerem que a transitividade não seria propriedade dos verbos, mas antes dos próprios contextos, ou de verbos em determinados contextos. Mas isso tem como conseqüência o esvaziamento da noção de transitividade, que se torna supérflua; e, como essa posição é comum, acho interessante criticá-la aqui com algum detalhe.

Digamos que, em vez de definir o verbo *comer* como transitivo ou intransitivo, disséssemos que ele é transitivo (ou "usado transitivamente") em (6) e intransitivo (ou "usado intransitivamente") em (7) e (14). As categorias "transitivo" e "intransitivo" já não se poderiam aplicar ao verbo *comer* tal como se apresenta no léxico, isto é, fora de contexto. *Comer* seria transitivo quando ocorresse com objeto direto, e intransitivo quando ocorresse sem objeto direto.

Mas isso viola a definição tradicional de transitividade: com efeito, segundo a definição, um verbo é transitivo não quando ocorre com OD, mas quando **exige** a presença de um OD — ou seja, quando **sempre** aparece com OD. Na verdade, essa tentativa de solucionar o problema equivale a estabelecer sinonímia entre **ser transitivo** e **ocorrer com OD**; por conseguinte, a noção de

"transitivo" deixaria de ser útil, pois não faria mais que repetir a informação já dada pela expressão menos misteriosa "que tem objeto direto".

É necessário distinguir dois tipos de informação sobre os itens léxicos, a saber: (a) em que contexto o item ocorre em uma frase dada (relação sintagmática); e (b) em que contextos o item **pode** ocorrer (relação paradigmática). A informação tipo (a) é particularizada e é fornecida pela análise da estrutura em questão. Por exemplo, no contexto da frase (6), *meu gato* é sujeito; *comer* ocorre com objeto direto etc. Já a informação tipo (b) é generalizada e é fornecida para o item em "estado de dicionário", fora de contexto — muito embora só possa ser depreendida, evidentemente, a partir do exame dos contextos. Assim, *meu gato* pode ser sujeito e pode também ser objeto direto, mas não pode ser núcleo do predicado, nem adjunto oracional. *Comer* pode ser núcleo do predicado, e não modificador; e pode ocorrer com ou sem objeto direto.

Ambos os tipos de informação são fundamentais para a tarefa de descrição da língua. Se não tivermos informações sintagmáticas [tipo (a)], a descrição não terá base nos fatos. Mas, se renunciarmos à informação paradigmática [tipo (b)], estaremos desistindo de expressar em termos gerais o comportamento gramatical das unidades lingüísticas. Um pouco de reflexão deve bastar para deixar claro que as conseqüências seriam destrutivas para a gramática; por exemplo, inviabilizaria o estabelecimento de classes de palavras. Temos de concluir que tentar definir "transitividade" em termos de contextos específicos não cura os males da concepção tradicional.

Voltemos à definição de "transitividade" como **exigência** de objeto direto, e "intransitividade" como **recusa** de objeto direto. Há verbos que efetivamente se comportam dessa maneira. Assim, *fazer* só ocorre com objeto direto:

(15) a. Evaristo faz lindas cortinas.
b. * Evaristo faz.

Desse modo, *fazer* seria "transitivo" segundo a definição tradicional, e essa classificação não apresenta problemas, já que esse verbo só ocorre com OD.

O verbo *nascer* só ocorre sem OD:

(16) a. Meu irmãozinho nasceu no sábado.
b. * Meu irmãozinho nasceu um nascimento tranqüilo.

logo, seria "intransitivo".

O problema surge quando verificamos que há verbos, como *comer*, que podem ocorrer com ou sem OD, sem por isso causar inaceitabilidade. Vimos que não há lugar para tais verbos no sistema tradicional; uma análise adequada deve, primeiramente, criar esse lugar.

6.2.2. Desenvolvendo uma nova análise

É necessário reconhecer que as duas noções de "exigência" e "recusa" de complementos não são suficientes para caracterizar todos os tipos de transitividade encontrados na língua. Vamos acrescentar uma terceira, a de "aceitação livre".

Agora, poderemos dizer que o verbo *nascer* **recusa** objeto direto; o verbo *fazer* **exige** objeto direto; e o verbo *comer* **aceita livremente** objeto direto. Representamos essa informação como traços dos respectivos verbos, da seguinte maneira:

> *nascer* tem o traço [Rec-OD] ("recusa OD");
>
> *fazer* tem o traço [Ex-OD] ("exige OD");
>
> *comer* tem o traço [L-OD] ("aceita livremente OD").

Como se vê, as duas classes tradicionais de "transitivos" e "intransitivos" se desdobram em pelo menos três classes de verbos, a saber, os que recusam OD (marcados [Rec-OD]), os que o exigem ([Ex-OD]) e os que o aceitam livremente ([L-OD]). Desse modo, sanamos o primeiro, e básico, defeito da análise tradicional da transitividade verbal. A descrição das transitividades deve ser feita em termos de **exigência**, **recusa** e **aceitação livre** de cada uma das funções relevantes.

Isso nos leva à segunda questão a solucionar. Segundo a análise tradicional, algumas funções são relevantes para o estabelecimento das transitividades, outras não. Assim, os verbos são transitivos ou intransitivos segundo exijam ou recusem objeto; e os transitivos são diretos, indiretos ou diretos e indiretos segundo o objeto exigido seja direto, indireto ou ambos, respectivamente. Os verbos de ligação exigiriam predicativo do sujeito. Dessa forma, ficam selecionadas algumas funções como relevantes, a saber, objeto direto, objeto indireto e predicativo do sujeito. Não há verbos que "exijam" ou "recusem" sujeito, por exemplo, ou adjunto adverbial etc.

A mesma seleção de funções precisa ser feita em termos de nossa análise atual, mesmo porque as próprias funções foram reformuladas na seção 3.2. Passamos agora, portanto, a considerar a pergunta: Que funções sintáticas são relevantes para o estabelecimento das transitividades?

As funções relevantes são aquelas que são **exigidas** ou então **recusadas** por algum verbo. Uma função que seja aceita livremente por todos os verbos não é relevante para estabelecer a transitividade, pois não caracteriza os verbos com que coocorre.

Não é difícil ver por que é assim. Tomemos uma função livremente aceita por qualquer verbo: o adjunto oracional (AO). Observa-se que, qualquer que seja o verbo de uma oração, é sempre possível colo-

car (ou deixar de colocar) nessa oração um adjunto oracional. Se marcarmos cada verbo que aceita AO como [L-AO], na verdade estaremos perdendo nosso tempo; é muito mais simples estabelecer, de uma vez por todas, que todos os verbos da língua aceitam AO. Nesse caso, o AO nem sequer precisa ser mencionado ao se exprimir a transitividade de cada verbo.

A situação é diferente com o OD: alguns verbos o exigem, outros o recusam, e outros o aceitam livremente. Nesse caso, é necessário especificar, para cada verbo, qual é seu comportamento quanto à possibilidade de ocorrência do OD; nenhuma afirmação pode ser feita de uma vez por todas, em geral, para todos os verbos da língua. Concluímos que o OD é uma função relevante para a formulação das transitividades.

As pesquisas realizadas recentemente sobre a transitividade mostram que há quatro funções relevantes em português, a saber: **objeto direto**, **complemento do predicado**, **predicativo** e **adjunto circunstancial**. Todas as outras funções são irrelevantes, por serem aceitas livremente por qualquer verbo. Não parece haver caso de função exigida por todos os verbos; se houvesse, essa função também seria irrelevante, por motivos óbvios.

A única dúvida a respeito da lista de funções relevantes deriva da transitividade do verbo *haver* "existir", caso este deva ser considerado um item léxico diferente de *haver* "auxiliar". Nesse caso, *haver* seria um verbo (o único da língua) que recusa sujeito. Como está em aberto a questão de se os dois verbos *haver* são de fato dois itens ou um único item, vamos ignorar esse caso aqui. Cada verbo, portanto, deverá ser marcado quanto a sua transitividade para cada uma das quatro funções relevantes: **OD**, **CP**, **Pv** e **AC**.

O levantamento feito por Barros (1992; 1993) mostra que as duas principais situações reconhecidas pela classificação tradicional, a saber, exigência de OD ([Ex-OD]) e recusa de OD ([Rec-OD]), são minoritárias na língua; a maior parte dos verbos é marcada [L-OD], não tendo assim lugar na classificação tradicional. Tendo examinado cerca de 550 verbos, Barros apurou que 58,9% deles são marcados [L-OD]; 31,5% são marcados [Ex-OD]; e apenas 9,6% são marcados [Rec-OD]. Esses números mostram que a classificação tradicional deixa escapar justamente o grupo mais importante de verbos: os que aceitam livremente a ocorrência de objeto direto.

A transitividade completa de um verbo é representada por um grupo de traços, um para cada função relevante. Assim, a transitividade de *comer* é [L-OD, Rec-CP, L-AC, Rec-Pv]: esse verbo aceita livremente o OD e o AC, e recusa o complemento do predicado e o predicativo. Essa é a matriz mais comum na língua e vale para quase todos os verbos marca-

dos [L-OD] — exatamente 57,6% de todos os verbos. A seguir, dou a lista de todas as matrizes de transitividade existentes em português, juntamente com a porcentagem de verbos que seguem cada uma, e um exemplo de cada:

Matrizes de transitividade verbal

I. [L-OD, L-AC, Rec-Pv, Rec-CP] (57,6%): *comer*

II. [Ex-OD, L-AC, Rec-Pv, Rec-CP] (22,3%): *encontrar*

III. [Rec-OD, L-AC, Rec-Pv, Rec-CP] (5,1%): *acontecer*

IV. [Rec-OD, Ex-AC, Rec-Pv, Rec-CP] (3,7%): *morar*

V. [Ex-OD, Ex-AC, Rec-Pv, Rec-CP] (2,1%): *acostumar*

VI. [Ex-OD, L-AC, L-Pv, Rec-CP] (1,3%): *considerar*

VII. [L-OD, L-AC, L-Pv, L-CP] (0,7%): *julgar*

VIII. [L-OD, L-AC, Rec-Pv, L-CP] (0,6%): *permanecer*

Além dessas seis matrizes, há outras três, para certos verbos que apresentam a peculiaridade de exigir **OD ou AC**, ou então **CP ou AC**, ou ainda **CP ou Pv**. Por exemplo, os verbos que seguem a primeira dessas matrizes podem ocorrer sem OD, mas nesse caso devem ter AC; ou sem AC, mas nesse caso precisam ter OD; e assim paralelamente para as outras duas matrizes.

IX. [Ex-(OD v AC), Rec-Pv, Rec-CP] (5,2%): *lembrar*

X. [Ex-(CP v AC), Rec-OD, Rec-Pv] (0,7%): *estar*

XI. [Ex-(CP v Pv), Ex-OD, L-AC] (0,7%): *sentir*

O exemplo mais conhecido de verbo que faz exigência alternativa é *estar*, que ocorre com CP, como em

(17) Mariela está adoentada.

ou com AC, como em

(18) Mariela está na cozinha.

mas nunca sem nenhum dos dois:

(19) * Mariela está.

A lista de matrizes dada acima delineia uma situação radicalmente diferente da que nos é fornecida pela gramática tradicional. Segundo a gramática tradicional, somente as relações "recusa" e "exigência" é que valem para estabelecer transitividades; aqui, acrescentamos "aceitação livre". A gramática tradicional considera relevantes as funções de objeto direto, objeto indireto e predicativo do sujeito; para nós, são relevantes o objeto direto, o adjunto circunstancial (que inclui os casos tradicionais de "objeto indireto", mais muitos outros casos), o complemento do predicado (correspondente aproximado do "predicativo do sujeito") e o predicativo (que corresponde, aproximadamente, ao "predicativo do objeto" tradicional). O resultado é que tradicionalmente se distinguem cinco subclasses de verbos (transitivos diretos, transiti-

vos indiretos, transitivos diretos e indiretos, intransitivos e de ligação), ao passo que em nosso sistema há onze (correspondentes às onze matrizes vistas acima).

Mas a diferença decisiva é a seguinte: as onze matrizes aqui propostas e definidas dão conta de descrever a transitividade de todos os verbos da língua (com a possível exceção, já mencionada, de *haver*), ao passo que as cinco subclasses tradicionais deixam de fora a maioria dos verbos, a saber: 58,9% que aceitam livremente o OD, além de muitos verbos que exigem AC em casos em que o AC não seria analisado como objeto indireto pela gramática tradicional.

Além disso, as onze matrizes se baseiam em funções cujas definições são puramente sintáticas e muito mais facilmente operacionalizáveis do que as definições tradicionais (no que pesem as confessadas incertezas que permanecem no atual sistema). Por tudo isso, não me parece caber dúvida de que o sistema acima exposto oferece um quadro muito mais adequado do fenômeno da transitividade verbal em português do que qualquer outro atualmente disponível.

Esta análise do sistema de transitividade do português, é preciso reconhecer, é mais complexa do que a tradicional. Temos quatro funções relevantes (OD, AC, CP, Pv) e, para cada uma delas, três possibilidades (Ex, Rec, L), o que nos dá dezenas de maneiras possíveis de exprimir a transitividade de um verbo — em contraste com as cinco maneiras tradicionais. Ainda que não houvesse maneira de simplificar esse sistema, a complicação não seria gratuita, pois se os fatos são complexos não há meio de descrevê-los de forma simples. A nova análise deve ser defendida em termos de sua capacidade superior de representar os fatos da língua.

No entanto, a complicação não é tão grande quanto parece, porque das matrizes possíveis em princípio somente onze se realizam; a grande maioria das possibilidades não se aplica a nenhum verbo e fica excluída da gramática. Falta fazer uma restrição importante à aplicação da tabela de transitividades dada acima. A tabela estipula, por exemplo, que o verbo *encontrar* exige objeto direto. Mas, naturalmente, é possível construir uma oração bem formada com *encontrar* e sem objeto direto: basta construir uma oração passiva, pois não pode ocorrer OD nas passivas. Seria o caso de incluir essa restrição na tabela?

Na verdade, isso não é necessário: representaria uma complicação considerável (pois o problema não se limita às passivas), e há uma maneira mais simples de fazê-lo. Basta estabelecer uma convenção geral: as transitividades expressas na tabela valem apenas para **estruturas não parcialmente correspondentes.** As ocorrências dos verbos em estruturas parcialmente correspondentes

podem ser deduzidas das próprias fórmulas de correspondência, mais os traços de correspondência dos verbos. As passivas (como se viu no capítulo 2) são parcialmente correspondentes das ativas. E cada verbo deve ser marcado quanto a poder ou não ocorrer em estruturas passivas (esse tipo de marca, chamada **traço de correspondência**, ocorre em todos os casos de correspondência parcial). Ora, um verbo que seja marcado como podendo ocorrer em passivas, seja qual for sua classe de transitividade, pode ocorrer sem OD — isso em uma estrutura passiva, que, sendo parcialmente correspondente, não conta para o estabelecimento das classes de transitividade.

Assim, o verbo *encontrar* pode continuar sendo marcado [Ex-OD]; isso não significa propriamente que ele **nunca** ocorre sem OD; mas significa que nunca ocorre sem OD, a não ser em estruturas parcialmente correspondentes, que são assim consideradas como que derivações das estruturas básicas.

6.2.3. Sintaxe ou semântica?

Há ainda uma diferença muito importante entre a concepção de transitividade aqui apresentada e a que se encontra nas gramáticas usuais. A idéia tradicional de transitividade é predominantemente semântica; procura-se justificar exigências e recusas em termos do significado de cada verbo. Uma passagem típica é a seguinte:

> [os verbos transitivos], para poderem formar o predicado, têm necessidade de um complemento (= objeto) que lhes integra o significado. [Kury, 1972, p. 31]

Em outro local, o mesmo autor afirma que a predicação (= transitividade) é um "fato sintático" (p. 30); mas creio que para ele o termo "sintático" não exclui fatores ligados ao significado, de modo que não há necessariamente contradição entre essas duas passagens. Mas, se delinearmos "sintaxe" e "semântica" da maneira como se faz neste trabalho (e na lingüística moderna em geral), teremos de dizer que a maioria dos gramáticos coloca os fenômenos de transitividade no domínio da semântica. Isso equivale à hipótese de que as possibilidades de ocorrência dos diversos complementos de um verbo seriam previsíveis, pelo menos em grande parte, a partir da semântica desse verbo. Conseqüentemente, seria redundante formular essas possibilidades dentro da sintaxe.

A concepção de transitividade aqui adotada é puramente sintática: lança mão das funções sintáticas "objeto direto", "adjunto circunstancial", "complemento do predicado" e "predicativo", que foram definidas formalmente em 3.2., e marca cada verbo sem referir-se a traços de seu significado. No entanto, isso não significa que a transitividade de um verbo não

possua correlato semântico algum; pode-se argumentar que, se *comer* admite objeto direto, isso é derivado do fato de que exprime uma ação que envolve um paciente (uma "coisa comida"). Eu colocaria a questão nos seguintes termos: não há dúvida de que existe certo grau de correlação entre, de um lado, as exigências e recusas feitas pelos verbos quanto à ocorrência de complementos e, de outro lado, os traços semânticos desses verbos. Mas essa correlação não é tão direta e tão generalizada a ponto de se poder prever a transitividade de um verbo a partir de sua semântica.

Em favor dessa posição podem-se aduzir alguns argumentos. Observe-se, para começar, que há casos de grupos de verbos cuja semântica é quase a mesma, mas que diferem quanto à transitividade: *morrer* e *falecer* são um bom exemplo. Sabe-se que *morrer* pode ocorrer com objeto direto (chamado "objeto interno" por ser representado por uma nominalização do próprio verbo). Assim, temos

(20) Machado morreu uma morte tranqüila.

Já *falecer* e *perecer*, embora de significado muito próximo do de *morrer*, não admitem objeto direto em nenhuma circunstância:

(21) * Machado faleceu um falecimento tranqüilo.
(22) * Machado pereceu uma morte trágica.

Dessa maneira, somos obrigados a atribuir aos três verbos traços de transitividade diferentes: *morrer* é marcado [L-OD], mas *falecer* e *perecer* são marcados [Rec-OD]. Isso dificilmente poderia ser previsto a partir da semântica dos três verbos, que é quase a mesma.

Um segundo argumento deriva da observação de que as exigências semânticas dos verbos têm reflexos variados nas exigências de complementos formais. Por exemplo, *roubar* pode ocorrer com OD exprimindo a coisa roubada, como em

(23) Roberto roubou muito dinheiro da União.

mas também pode ocorrer com OD exprimindo a entidade que sofreu o roubo, como em

(24) Roberto roubou a União em muito dinheiro.

Como derivar essas duas construções da semântica, se as duas frases são sinônimas?

Um argumento semelhante pode ser construído a partir de frases como

(25) Maria espancou Mané.

em que o sujeito pratica a ação e o OD a recebe, e frases como

(26) Mané apanhou de Maria.

em que, apesar de a ação ser a mesma, o sujeito a recebe e o AC a pratica.

Em muitos casos, a interpretação inclui necessariamente uma função semântica, mas a expressão desta por um complemento é ou não obrigatória, segundo o verbo empregado. Por exemplo, *colocar* é sempre interpretado como tendo um paciente; ou seja, sem a "coisa colocada" não se pode ter a "ação de colocar". Esse paciente é obrigatoriamente expresso por um objeto direto (ou por um sujeito em frases passivas):

(27) Elisa colocou o peixinho na geladeira.

Por outro lado, a idéia de "comer" igualmente exige um paciente, caso contrário, não poderá ter havido a "ação de comer"; mas a expressão desse paciente com o verbo *comer* é opcional:

(6) Meu gato já comeu todo o mingau.
(7) Meu gato já comeu.

É claro que, mesmo em frases como (7), entende-se necessariamente que o gato comeu alguma coisa. Mas ali não há **objeto direto** nenhum, ao contrário de (6) e (27). Como se vê, ainda aqui as exigências da transitividade não refletem as exigências do significado. Em vista desses fatos, fica difícil derivar a transitividade dos verbos a partir de sua semântica, ou vice-versa. Do ponto de vista das exigências semânticas, dois verbos como *colocar* e *comer* são idênticos porque ambos supõem a existência de um paciente. Mas esses verbos são diferentes no que diz respeito a suas exigências de complemento: *colocar* exige objeto direto, *comer* não exige. Aqui, como em tantos outros pontos, evidencia-se a necessidade de separar o aspecto semântico e o formal da descrição. Uma noção como "paciente" é um dos ingredientes do significado; já "objeto direto" é um dos ingredientes da organização formal das orações. Embora as noções sejam ambas indispensáveis, não se confundem em nenhum momento.

Concluí-se que é preciso incluir na descrição gramatical a transitividade como fenômeno formal (sintático). Cada verbo estabelece suas exigências quanto à ocorrência de complementos, e essas exigências são, em princípio, independentes de seu significado. Evidentemente, existe uma relação entre transitividade e traços do significado dos verbos; e muitos desses traços se referem a exigências como de "agente", "paciente" etc. Apenas, como costuma acontecer, a relação entre os dois planos de análise é complexa — um tema interessante de pesquisa, e não uma correlação simples que possa ser postulada de início.

6.2.4. Previsão de ocorrências

Há ainda uma pergunta a ser respondida a respeito da transitividade verbal: a transitividade será suficiente para prever quando um verbo

aparecerá com ou sem algum complemento, ou para explicar por que ocorreu ou não? Creio que algumas pessoas tenderiam a responder afirmativamente: quando temos um "verbo transitivo", sempre temos um objeto direto na mesma oração. No entanto, as coisas não são bem assim, conforme tentarei mostrar.

Tomemos um verbo marcado [Ex-OD], como *colocar*, por exemplo. O traço sugere que esse verbo só aparece com OD, e isso é verdade, exceto em condições de anáfora, como em

(28) Elisa pegou o peixinho e colocou na geladeira.

Nesse tipo de construção (aliás, muito típica da língua portuguesa), o segundo verbo pode ocorrer sem OD porque se segue a uma oração cujo OD é idêntico; interpreta-se que Elisa colocou o peixinho na geladeira (ver o capítulo 11 para maiores detalhes da interpretação de tais estruturas).

Geralmente frases anaforicamente reduzidas como (28) são desprezadas para efeito da formulação das transitividades; há boas razões em favor dessa solução, de modo que vamos aceitá-la aqui. Assim, poderemos dizer que *colocar*, assim como todos os verbos marcados [Ex-OD], realmente só ocorrem com objeto direto, nunca sem ele. Como se vê, em casos de **exigência**, a transitividade é suficiente para explicar a ocorrência de complementos com o verbo.

Em casos de recusa, dá-se o mesmo: o verbo *perecer* é marcado [Rec-OD] e, com efeito, nunca aparecerá com objeto direto.

Mas os casos de ocorrência livre apresentam um problema: como prever se um verbo como *comer*, marcado [L-OD], aparecerá com ou sem objeto direto em determinada frase? A própria natureza do traço impossibilita essa previsão. Em casos de ocorrência livre, a transitividade não basta para prever a ocorrência do complemento; *comer* aparecerá com ou sem objeto direto, dependendo de fatores extra-sintáticos, tais como: o significado da sentença, o grau de especificidade que o falante deseja dar à sua mensagem, os conhecimentos que o falante supõe que seu interlocutor já possua etc. A previsão, pois, só será possível se estivermos de posse de uma série de conhecimentos de diversa natureza: informação sintática (a transitividade do verbo) + informação semântica (o significado da frase) + conhecimentos acerca da situação de comunicação em que a frase é usada.

Em resumo: a transitividade só é suficiente para prever a ocorrência de um complemento em casos de recusa ou exigência; em casos de aceitação livre, a transitividade não fornece base suficiente para a previsão.

A esperança de prever a ocorrência dos complementos, em todos os casos, com base apenas na transiti-

vidade talvez tenha sido a causa de se definir transitividade como propriedade do contexto (ver 6.2.1.) — uma definição que, como vimos, destrói toda a relevância descritiva da noção de "transitividade".

Para deixar a questão um pouco mais clara, vou dar alguns exemplos de situações em que somente fatores extra-sintáticos é que permitem explicar **por que** determinado complemento ocorreu ou deixou de ocorrer. Vejamos primeiramente as frases

(29) André cuspiu em cima de mim.

(30) André cuspiu guaraná em cima de mim.

A primeira coisa a observar é que *cuspir* deve ser marcado [L-OD]. Isso explica por que pode ocorrer com OD, como em (30), ou sem ele, como em (29). Mas qual será a razão de ter ocorrido com OD em (30)?

A resposta (bastante óbvia) é que, em (30), afirma-se de André que cuspiu alguma coisa que não é o que normalmente se cospe. Como a "coisa cuspida" é geralmente cuspe, ela não precisa necessariamente ser expressa, embora possa sê-lo. Assim, entendemos (29) como afirmando que André lançou cuspe (e não outra coisa qualquer) em cima de mim. Aqui está em funcionamento um princípio de economia que estabelece que não se deve dizer mais (nem menos) do que o necessário. Com o verbo *cuspir*, não é necessário exprimir a coisa cuspida, quando esta é cuspe; mas, se for outra coisa qualquer, ela deve ser expressa.

Observe-se como a explicação de por que *cuspir* ocorre sem OD em (29) precisa lançar mão de conhecimentos de vários níveis: primeiro, sintáticos — *cuspir* é marcado [L-OD]. Depois, semânticos: a "coisa cuspida" é cuspe. Além disso, conhecimento do mundo em que vivemos: em geral, o que se cospe é cuspe. Finalmente, conhecimento de regras da comunicação (chamadas "máximas da conversação"): no caso, a máxima de que, quando uma informação é altamente previsível, não é necessário explicitá-la. Sabemos que a sintaxe é apenas um dos componentes da fala; a explicação cabal dos fenômenos observados precisa lançar mão de todos os componentes, lingüísticos e não-lingüísticos. A tentativa tradicional de explicar ocorrências com base apenas na transitividade é, portanto, um empreendimento sem esperança de sucesso.

Outros casos podem ser explicados seguindo essas mesmas linhas. Embora o verbo *pôr* ocorra quase sempre com objeto direto, há certas situações em que ele pode ocorrer sem objeto direto, como em

(31) Essa galinha põe todos os dias.

Pôr sem OD só aparece em situações muito particulares, impossíveis de caracterizar gramaticalmente. Trata-se de uma especialização do verbo para determinada situação

(no caso, de aves botando ovos), que pode ser considerada gramaticalmente marginal e não levada em conta para efeitos do estudo dos traços sintáticos do verbo.

Finalmente, vale mencionar o efeito do contexto situacional. Assim, o verbo *fazer* é marcado [Ex-OD], mas pode aparecer sem objeto direto se a situação de comunicação deixar suficientemente claro qual é a "coisa feita". Digamos que um grupo de crianças está fazendo bonecos de massa; aproxima-se uma outra criança e diz

(32) Eu também vou fazer!

A situação sendo suficiente para esclarecer de que se trata, o objeto direto não precisa ser expresso.

Todos esses exemplos mostram como a sintaxe se integra com os outros componentes da comunicação lingüística, para produzir mensagens eficientes. E mostram também como é às vezes difícil a tarefa do sintaticista, obrigado a separar, para estudo, componentes que na natureza aparecem integrados de maneira complexa.

6.3. TRANSITIVIDADE NOMINAL

Ainda não foi realizado, que eu saiba, um estudo detalhado da transitividade nominal, comparável aos que existem para a transitividade verbal. Assim, não será possível oferecer generalizações e levantamentos estatísticos como os que acabamos de ver para os verbos. Vou, portanto, apenas definir e exemplificar o fenômeno.

A análise tradicional não limita a transitividade aos verbos; considera-se que também certas palavras de outras classes — substantivos, adjetivos e advérbios — podem exigir, ou recusar, a presença de certos termos. Esses termos se analisam tradicionalmente como "complementos nominais"; para nós, serão modificadores, complementos do sintagma adjetivo ou complementos do sintagma adverbial, segundo o caso. Exemplos clássicos são:

crença em duendes;
favorável ao réu;
favoravelmente ao réu.

A idéia tradicional é que o complemento "completa o sentido" do substantivo, adjetivo ou advérbio da mesma forma que um objeto, por exemplo, "completa o sentido" de um verbo. Fala-se, por isso, de "palavras de predicação incompleta", que exigiriam um complemento para que seu significado fosse completo.

Essa análise é passível das mesmas objeções feitas a respeito dos complementos verbais. Na maioria dos casos, as palavras ditas "de predicação incompleta" podem perfeitamente aparecer sem complemento, em situações não-anafóricas, o que mostra que não "exigem" complementação:

(33) Você precisa respeitar minhas crenças.

(34) O tempo se mostrava favorável.

Analiso o termo *em duendes* no sintagma nominal *crença em duendes* como modificador, e portanto não faço a distinção tradicional entre "adjunto adnominal" e "complemento nominal". A razão é que a única diferença coerentemente formulada que encontrei nas gramáticas entre essas duas funções é que o complemento nominal é (ou pode ser) exigido por algum item, ao passo que o adjunto adnominal nunca o é. Mas isso não pode servir de base para uma distinção entre funções, pois a expressão da transitividade depende, ela própria, da definição prévia das funções.

Desse modo, o problema se planteia em nossos termos da seguinte forma: o fenômeno da transitividade nominal existirá, caso existam substantivos que exijam ou recusem modificador; adjetivos que exijam ou recusem CSA; e/ou outros tipos de palavras ("advérbios") que exijam ou recusem um complemento semelhante. Se não encontrarmos tais exigências ou recusas de complemento, a transitividade nominal ficará esvaziada, pois todas as funções serão irrelevantes, e teremos de dizer que tal fenômeno não existe na língua. Não parece ser esse o caso. Há, por exemplo, certo número de adjetivos que exigem CSA. Assim, os sintagmas seguintes são mal formados se não receberem complemento:

(35) * Um cliente desejoso

(36) * Um documento alusivo

(37) * Uma linha de ação tendente

Esses sintagmas se tornarão aceitáveis se acrescentarmos complementos adequados, como, respectivamente, *de atenção, à Independência do Brasil* e *à renúncia*. Por conseguinte, existem exigências de CSA feitas por adjetivos, e a transitividade nominal é um fato do português.

Há igualmente adjetivos que recusam CSA. Um desses é *presidencial*, que pode vir acompanhado de intensificador em certos casos, mas nunca de CSA. Isso confirma a necessidade de incluir a transitividade nominal na gramática.

Já com os substantivos os fatos são menos claros e estão a pedir mais investigação. Faltam, no momento, exemplos realmente convincentes de recusa ou exigência de modificador. Os possíveis exemplos encontrados (todos de exigência) são aceitos sem complemento pelo menos por alguns falantes:

(38) Um partido cheio de tendências (à desagregação)

(39) Esse documento contém alusões (à Independência)

(40) Todos se assustaram com o aparecimento (do fantasma)

Ao se omitir o modificador, o resultado é muito melhor do que qualquer dos exemplos de (35) a (38).

Como se vê, é necessário deixar parcialmente em suspenso a questão do modo como se manifesta a transitividade nominal em portu-

guês e em que dimensões; faltam pesquisas relevantes a respeito. O que se pode adiantar desde já é que a transitividade nominal é um fato da língua; aplica-se aos adjetivos e possivelmente também aos substantivos e advérbios. A alternativa normal parece ser a aceitação livre ("L"); recusas e exigências são comparativamente excepcionais.

6.4. REGÊNCIA: FORMA DOS COMPLEMENTOS ORACIONAIS

6.4.1. Subjuntivo e indicativo

Um aspecto importante do fenômeno da regência são as exigências feitas pelos verbos, e também por outros itens, quanto à forma do verbo de seus complementos oracionais; em particular, essas exigências se referem ao **modo** do verbo do complemento. Nos exemplos abaixo,

(41) Lelé demonstrou que podia fazer o serviço.

(42) Lelé duvidou que pudesse fazer o serviço.

observa-se uma diferença de modo no verbo da oração subordinada: está no indicativo em (41) e no subjuntivo em (42). Conforme veremos, essa diferença pode ser atribuída ao verbo da oração principal.

Observa-se um fenômeno paralelo com orações subordinadas a outros itens que não verbos. Assim, uma oração subordinada à preposição *para* aparece no subjuntivo, mas não no indicativo:

(43) Trouxemos este frango para que você o mate.

Já, com a preposição *até*, ambos os modos são aceitáveis, com diferença semântica, no caso:

(44) Fiquei escondido até que você chegou.

(45) Ficarei escondido até que você chegue.

Há ainda palavras consideradas "advérbios" que regem o modo da oração a que pertencem. *Talvez* aceita subjuntivo quando o verbo está à sua direita, como em

(46) a. Eu talvez o procure no escritório.

Em geral, aceita-se também o indicativo nesses casos, mas me parece que é mais raro:

b. Eu talvez o procurarei no escritório.

Mas, se o verbo está à esquerda de *talvez*, só o indicativo é admissível:

(47) a. Eu o procurarei no escritório, talvez.

b. * Eu o procure no escritório, talvez.

Outro fator que determina o aparecimento do subjuntivo em certos casos é a presença de uma negação verbal. Por exemplo, com o verbo *afirmar*, usa-se o indicativo:

(48) a. O advogado afirmou que você é inocente.
b. * O advogado afirmou que você seja inocente.

Mas, se há uma negação verbal na principal, a subordinada pode vir no subjuntivo (assim como no indicativo):

(49) a. O advogado não afirmou que você é inocente.
b. O advogado não afirmou que você seja inocente.

Finalmente, em certos casos, parece que o tempo verbal pode condicionar a ocorrência do modo da subordinada. Tenho em mente exemplos como

(50) a. Ele pensou que estivesse/estava protegido.
b. Ele pensava que estivesse/estava protegido.
c. Ele pensa que * esteja/está protegido.

Como se vê, em (c) apenas o indicativo dá resultado aceitável, o que se correlaciona com o fato de que o verbo da principal está no presente. Em muitos dos casos em que os dois modos são possíveis, percebe-se uma diferença de significado mais ou menos nítida entre as duas versões. Isso pode ser atribuído à própria semântica da forma subjuntiva.

As condições que governam a ocorrência do subjuntivo, em oposição ao indicativo, em todas essas construções são complexas, e bastante mal compreendidas, apesar de muita pesquisa já realizada na área. Nesta seção, apresentarei alguns resultados que, embora parciais, deverão dar uma idéia do funcionamento desse complicado mecanismo. A exposição só abordará o caso de complementos oracionais em função de **objeto**, porque para esses há disponibilidade de estudos mais completos; falta ainda um estudo abrangente da forma dos complementos oracionais em função de sujeito, por exemplo.

6.4.2. Infinitivo

Nos casos vistos acima, a subordinada é sempre introduzida por um elemento (em geral *que*, mas às vezes também *se*), denominado **complementizador**. Há ainda subordinadas sem complementizador explícito: o verbo nesses casos fica no infinitivo ou no gerúndio. Vamos considerar aqui apenas o infinitivo.

A ocorrência de subordinadas no infinitivo é em parte regida por condições semelhantes às que regem o subjuntivo ou o indicativo. Assim, há verbos que aceitam complemento no infinitivo:

(51) Lelé demonstrou estar em boas condições físicas.

e há outros (bem menos numerosos) que não o aceitam:

(52) a. Lelé mentiu que estava em boas condições físicas.
b. * Lelé mentiu estar em boas condições físicas.

Também algumas preposições admitem infinitivo:

(53) Fiquei escondido até você chegar.

(54) Vou embora para você ficar mais feliz.

Essas preposições, como vimos nos exemplos (43) a (45), também podem ocorrer com o complementizador *que*. Algumas, entretanto, só aceitam infinitivo:

(55) a. Conversaremos após a aula acabar.
b. * Conversaremos após que a aula acabe.

Não se conhecem casos de advérbios que exijam infinitivo.

Na exposição que se segue, consideraremos as condições sob as quais os verbos se constroem com subjuntivo, indicativo ou infinitivo. A descrição é bastante preliminar, mas creio que cobre a maioria dos casos, podendo servir de base para um levantamento mais amplo.

6.4.3. Forma dos complementos oracionais do verbo

Aqui vamos estudar a forma dos complementos do verbo, exceto o sujeito. Em geral os exemplos serão com objeto direto oracional (isto é, que contém oração); mas em alguns casos podem surgir exemplos de outras funções.

As exigências feitas pelos verbos quanto à forma das orações que lhes são subordinadas são muito variadas, e aqui tentarei estudar as que me parecem mais importantes.

O complementizador é basicamente determinado pela forma do verbo: *que* só ocorre com indicativo ou subjuntivo, e só o infinitivo ocorre normalmente sem complementizador. Há casos de subjuntivo sem complementizador, mas podem ser considerados marginais; um exemplo é

(56) Solicito sejam deferidos estes pedidos.

Por outro lado, fatores importantes são: o modo da subordinada (subjuntivo, indicativo ou infinitivo); além disso, no caso de subordinadas no infinitivo, a possibilidade ou não de ocorrer sujeito, assim como a eventual exigência de preposição. Vejamos cada um desses fatores, com exemplificação.

Em primeiro lugar, o verbo pode exigir um modo específico. Alguns verbos exigem indicativo, como, por exemplo, *demonstrar*:

(41) Lelé demonstrou que podia fazer o serviço.

(57) * Lelé demonstrou que pudesse fazer o serviço.

Outros verbos exigem subjuntivo:

(42) Lelé duvidou que pudesse fazer o serviço.

(58) * Lelé duvidou que podia fazer o serviço.

Há ainda outros que aceitam ambos os modos (muitas vezes, como já apontei acima, com diferença de significado):

(59) a. Eu entendo que você quer ir embora.

b. Eu entendo que você queira ir embora.

Em português, a imensa maioria dos verbos aceita infinitivo, de uma maneira ou de outra. Em muitos casos, entretanto, há restrições; por exemplo, com alguns verbos o infinitivo só é aceitável se a subordinada não tiver sujeito:

(60) a. Manuel quer comprar o piano.

b. * Manuel quer Artur comprar o piano.

Já em outros casos, a presença do sujeito não impede a aceitabilidade:

(61) a. Lelé reconheceu estar em más condições físicas.

b. Lelé reconheceu o time estar em más condições físicas.

Além disso, alguns verbos só aceitam infinitivo se este estiver em um complemento preposicionado:

(62) a. Os alunos pediram para sair.

b. * Os alunos pediram sair.

Note-se que, com esse verbo, a preposição não é necessária se o complemento não estiver no infinitivo:

(63) Os alunos pediram que o professor saísse.

6.4.4. Classificação dos verbos

Desses fatos, é possível depreender cinco traços que caracterizam o comportamento dos verbos da língua quanto à forma dos complementos oracionais que aceitam. Estou deixando de lado outros traços de certa importância (como, por exemplo, a exigência de subjuntivo quando há negação verbal na principal), de forma que a classificação abaixo deve ser tomada como uma descrição parcial. De qualquer modo, acredito que temos condições de estabelecer uma classificação muito mais rica e informativa do que a tradicional.

Os traços que utilizarei na classificação dos verbos são os seguintes:

[Indic] – "admite indicativo na subordinada";

[Subj] – "admite subjuntivo na subordinada";

[Inf c/suj] – "admite infinitivo, com sujeito";

[Inf s/suj] – "admite infinitivo, sem sujeito";

[Inf prep] – "admite infinitivo apenas se houver preposição".

Note-se que vários fenômenos ficam de fora desse grupo de traços, notadamente a influência da negação [veja-se o exemplo (49)] e a do tempo do verbo da principal [veja-se o exemplo (50)].

Utilizando esses cinco traços, distinguem-se nove classes de verbos em português, conforme a tabela abaixo:

Classificação dos verbos quanto à forma do complemento

Classe	[Indic	Subj	Inf c/suj	Inf s/suj	Inf] prep	Exemplos
I	+	+	+	+	–	*admitir*
II	+	+	–	+	–	*entender*
III	+	–	+	+	–	*demonstrar*
IV	–	+	+	+	+	*pedir*
V	+	–	–	–	–	*mentir*
VI	–	+	–	–	–	*duvidar*
VII	–	+	–	+	–	*mandar, querer*
VIII	–	–	–	+	–	*dever*
IX	–	–	–	+	+	*atrever-se*

Notas

(a) A classe V é extremamente pequena: possivelmente, apenas o verbo *mentir.*

(b) *Mandar* não aceita infinitivo com sujeito porque em frases como *mandei Carolina lavar o carro* é possível mostrar que *Carolina* é objeto direto da oração principal.

(c) Os verbos da classe VIII são pouco numerosos: *dever, ousar, costumar, tramar* e alguns outros. Alguns são às vezes classificados como "auxiliares modais".

6.5. REVISITANDO OS PREDICADOS COMPLEXOS

Agora podemos voltar a examinar os predicados complexos vistos na seção 3.2.2.2., dando em maior detalhe as razões para sua análise. Vimos que em frases como

(64) Lili está lendo.

embora, de certo modo, tenhamos dois "verbos", *estar* e *ler,* há um só predicado. Esse é o único caso em que uma forma verbal não constitui um núcleo do predicado: o NdP é *lendo,* e *está* é o **auxiliar**. A seqüência Aux + NdP forma o predicado (que, neste caso, é complexo).

A razão principal para essa análise é o fato de que, para efeitos da maioria dos processos gramaticais, a seqüência *está lendo* em (64) funciona como um único verbo; na verdade, funciona como se apenas o verbo *ler* estivesse presente na oração; *está* é "transparente" a esses processos.

O principal deles é a transitividade. Observa-se que qualquer traço de transitividade que valha para um verbo quando pertence a um predicado simples vale igualmente para esse verbo quando é parte de um predicado complexo. Exprimimos

isso dizendo que **existe um conjunto de traços de transitividade para cada predicado**. Assim, por exemplo, o verbo *comer* tem a transitividade [L–OD, L–AC, Rec–Pv]. Ora, esses mesmos traços se aplicam a frases com os predicados *está comendo* ou *tem comido*. Isso se aplica a todos os verbos da língua; a conclusão é que não convém considerar dois predicados na seqüência *está comendo*, porque um deles seria excepcional por não fazer exigência nenhuma quanto à transitividade. Note-se que *estar*, por exemplo, quando sozinho em seu NdP, tem traços de transitividade próprios, expressos pela matriz [Ex–(CP v AC), Rec–Pv]. Mas essa matriz fica neutralizada quando *estar* ocorre em um predicado complexo (um argumento possível em favor de distinguir dois verbos *estar*). Assim, o verbo *considerar* aceita predicativo; e é possível construí-lo com predicativo mesmo quando vem acompanhado de *estar* (que recusa predicativo) em um predicado complexo:

(65) A chefe está considerando Carlão um bom funcionário.

Assim, os fatos relativos à transitividade apóiam a análise de seqüências do tipo *estar* + gerúndio e *ter* + particípio como predicados complexos.

Agora é possível dar uma razão de certa importância para que, em uma frase como (64), consideremos *lendo*, e não *está*, o núcleo do predicado. Dessa maneira, será possível atribuir a transitividade sempre ao núcleo do predicado, mantendo a generalização de que o NdP é o elemento "governante" da oração, para efeitos de transitividade. Esse foi o principal motivo que me levou a analisar a forma não-conjugada dos predicados complexos como NdP — mantendo, aliás, a intuição tradicional, que considera essa forma o "verbo principal".

6.6. CONCORDÂNCIA

6.6.1. Preliminares: traços do SN

Examinaremos agora outro fenômeno da língua que pode ser colocado sob o rótulo geral de "regência". Nos casos que vamos passar a estudar, vê-se tradicionalmente uma espécie de exigência de harmonização de flexões entre os diversos constituintes de uma construção. Chama-se a isso **concordância**. Como veremos nas seções seguintes, há dois tipos principais de concordância: a concordância entre o sujeito e o NdP de uma oração e a concordância entre diversos elementos nominais (tradicionalmente classificados como "substantivos", "adjetivos", "artigos", "numerais" e "pronomes").

Em todos esses casos, há a exigência de que certos traços (em geral expressos morfologicamente)

sejam idênticos em vários constituintes. Assim, em uma frase como

(66) Minhas sobrinhas ganharam um cavalo.

os traços de "terceira pessoa" e "plural", considerados como presentes no SN *minhas sobrinhas*, de certa forma se reproduzem no verbo que preenche o NdP: *ganharam*, e não, por exemplo, *ganhou*.

Outro exemplo é o SN *minhas sobrinhas*, onde *sobrinhas* é marcado no léxico como "feminino"; correspondentemente, o item *minhas* também fica no feminino: *minhas*, e não *meus*. Esses fenômenos serão estudados detalhadamente nas seções 6.6.2. e 6.6.3. Antes disso, teremos de examinar a maneira pela qual os constituintes adquirem os traços que serão relevantes para efeitos de concordância.

6.6.1.1. Marcação dos SNs: pessoa

Um dos tipos de concordância, como vimos, resulta em que um dos SNs da oração (o sujeito) apresenta alguns traços em comum com o NdP; isso já foi visto no capítulo 3, onde definimos o sujeito como o termo que está em relação de concordância com o NdP: [+ CV]. Vamos agora ver mais de perto o que significa isso.

Há um conjunto de formas verbais que são chamadas "de terceira pessoa": *chega, chegou, chegará, chegam, chegaram, chegarão* etc. Essas formas se harmonizam com a imensa maioria dos SNs: as três primeiras com SNs no singular, as três últimas com SNs no plural. Assim, temos:

(67) Minhas sobrinhas chegaram/chegam/chegarão de trem.

(68) Minha sobrinha chegou/chega/chegará de trem.

Nessas frases, o SN *minha(s) sobrinha(s)* está em relação de concordância com o NdP e é, portanto, o sujeito.

Mas há um pequeno número de itens que têm formas especiais do verbo para manifestar a concordância. Esses itens são considerados como "de primeira pessoa": *eu, nós*; ou "de segunda pessoa": *tu, vós*. Os itens de segunda pessoa raramente se usam no português padrão brasileiro de hoje, e portanto não os exemplificarei abaixo. Mas os de primeira pessoa são muito freqüentes, e com eles os verbos precisam assumir formas diferentes das que vimos nos exemplos (67) e (68):

(69) Eu cheguei/chego/chegarei de trem.

(70) Nós chegamos/chegamos/chegaremos de trem.

Considera-se, portanto, que esses itens são marcados, já no léxico, com traços de pessoa: *eu, nós* são marcados como de primeira pessoa; *tu, vós* como de segunda; e os outros SNs da língua, em geral, são

marcados como de terceira pessoa (ou, talvez melhor, não são marcados para pessoa).

Nos casos em que um SN é composto pela coordenação de constituintes de várias pessoas, aplica-se uma espécie de hierarquia, segundo a qual prevalece a pessoa de "menor número". Assim, um SN composto de elementos de primeira e de terceira pessoa é considerado como de primeira pessoa (e plural, cf. 6.6.1.2.):

(71) [Minhas sobrinhas e eu] chegamos de trem.

Note-se que o SN tem seu próprio traço de pessoa, embora este seja decorrente dos traços de pessoa de seus constituintes imediatos. Assim, diremos que o SN *minhas blusas e eu* tem o traço de primeira pessoa. Além disso, evidentemente, os constituintes também têm seus traços: *minhas blusas* é um SN marcado como de terceira pessoa, e *eu* é um SN marcado como de primeira. Finalmente, o item *blusas* é de terceira pessoa, e *eu* é de primeira. As regras dadas acima são suficientes para derivar, das marcas de pessoa das palavras que compõem um SN, o traço de pessoa do próprio SN. Os traços das palavras, relembro, são marcas léxicas, próprias a cada palavra individual. Para efeito de concordância com o NdP, vale o traço do SN maior de todos.

6.6.1.2. Marcação dos SNs: gênero e número

Além de traços de pessoa, os SNs também têm traços de gênero e número, e estes são relevantes para o segundo tipo de concordância mencionado acima, exemplificado no sintagma *minha blusa*. O item *blusa* é marcado no léxico como "feminino" e nisso se opõe, por exemplo, a *guarda-chuva*, que é "masculino". Esses traços exprimem apenas o comportamento desses itens no que se refere à concordância: *blusa* coocorre no SN com formas como *minha*, e *guarda-chuva* com formas como *meu*. Muitas palavras da língua são assim marcadas como "femininas" (*blusa, sobrinha, chuva, mão*) ou "masculinas" (*guarda-chuva, sobrinho, vento, pé*); essa marca, embora tenha certo grau de correlação com uma oposição de sexo, deve ser considerada aqui como puramente formal — descreve parte do comportamento morfossintático do item em questão, nada mais. A correlação com a noção (semântica) de "sexo" é muito imperfeita e indireta, e não precisamos ocupar-nos dela neste capítulo.

Outras palavras da língua não são propriamente marcadas no léxico como femininas ou masculinas; antes, **variam** em gênero, apresentando sistematicamente uma forma feminina e uma masculina: *meu/minha*; *novo/nova*; *branco/branca*. Em muitos casos as duas formas são fonologicamente idênticas, mas devem

ainda assim ser consideradas distintas; assim, *verde* é masculino em (72), feminino em (73):

(72) Um carro verde

(73) Uma janela verde

Em resumo, o gênero se manifesta de duas maneiras distintas: ou como propriedade inerente a um item léxico (*blusa* **é feminino**), ou como variação flexional (*meu* **está no masculino**, ou ainda **é o masculino de** ***minha***). Essa oposição recobre, *grosso modo*, a distinção tradicional entre "substantivos" e "adjetivos": os primeiros **possuem** gênero, os segundos **variam** em gênero.

Por outro lado, pode-se dizer que em sua maioria as palavras variam em número: em geral todas as palavras que possuem traços de número podem ser singulares ou plurais, conforme o caso: *meu/meus*; *blusa/blusas*. Algumas palavras só podem ocorrer no singular (*ouro*), e outras só no plural (*férias*), mas são excepcionais; a maioria esmagadora dos itens, sejam eles "substantivos", sejam "adjetivos", apresenta sistematicamente uma forma de singular e uma de plural. Ainda aqui, note-se, o singular e o plural podem ser idênticos (*lápis/lápis*), mas, tal como no caso do gênero, isso é considerado simples caso de coincidência fonológica.

Em geral, portanto, podemos dizer que as palavras **variam** em número, e não que **possuem** número.

A conseqüência é, pois, que as palavras das classes que nos interessam aqui sempre se manifestam como possuindo gênero (feminino ou masculino) e como estando em um número (singular ou plural). O gênero pode ser resultado de variação ou de marca léxica; o número, o mais das vezes, será resultado de variação.

Também aqui o SN tem seus traços próprios, derivados dos traços de seus componentes. A seguir, dou as regras que governam a atribuição dos traços de gênero e número ao SN.

(a) Quando não há divergência de gênero e número entre os constituintes imediatos de um SN, esses traços passam automaticamente ao SN.

Assim, o SN *minhas sobrinhas* é formado de dois itens marcados como "feminino", "plural"; o SN terá também os traços de feminino e singular.

Quando há divergência entre os constituintes imediatos do SN quanto a gênero e número, trata-se sempre de um caso de SNs coordenados formando um SN maior; todos os outros casos são mal formados e serão tratados na seção 6.6.3. No caso de SNs formados por coordenação, aplicam-se as seguintes regras:

(b) Todos os SNs formados por coordenação recebem o traço de número "plural".

(c) Quanto ao gênero, SNs formados por coordenação de SNs dos dois gêneros recebem o traço "masculino".

Assim, por exemplo, o SN *Antônio e Bia* é plural e masculino: plural por ser composto pela coordenação de dois SNs menores, o SN *Antônio* e o SN *Bia*; e masculino porque *Antônio* é masculino e *Bia* é feminino. Isso se evidencia na concordância:

(74) Antônio e Bia são casados.

Um SN composto apenas de femininos é feminino (e, naturalmente, plural):

(75) Marília e Bia são morenas.

6.6.1.3. Casos particulares e problemas

O mecanismo de atribuição de traços ao SN, exposto acima, nem sempre funciona com clareza. As gramáticas consignam certo número de casos particulares em que os princípios gerais não são seguidos. Alguns desses casos refletem usos arcaicos ou extremamente raros; mas outros merecem menção, por serem relativamente freqüentes. Relaciono adiante os mais importantes, mas seria necessário realizar levantamentos que nos dessem uma idéia clara da importância de cada caso na língua padrão atual.

O primeiro caso particular importante é o de SNs representados por relativos; como vimos na seção 5.3.3., tais SNs podem funcionar como sujeito (e, portanto, seus traços serão reproduzidos no NdP). Um exemplo é

(76) Só convidei os amigos que estavam aposentados.

A oração subordinada é *que estavam aposentados*; *que* é o sujeito. Como se pode ver pelo NdP e pelo CP, esse sujeito funciona como se tivesse os traços "masculino, plural, terceira pessoa"; mas a mesma palavra pode apresentar outros traços, como em

(77) Só convidei a amiga que estava aposentada.

Aqui os traços manifestados são "feminino, singular, terceira pessoa".

Evidentemente, o relativo toma seus traços do antecedente (respectivamente, *os amigos* e *a amiga*). Isso é obrigatório no caso dos relativos *que* e *o qual* — e este último também assume a forma correspondente aos traços adquiridos: *o qual, a qual, os quais, as quais.*

Já com o relativo *quem*, diz-se que, além de poder comportar-se como os demais, pode também assumir os traços "terceira pessoa, singular" mesmo se o antecedente não tiver aqueles traços. Esse fenômeno, a julgar pelos exemplos apresentados nas gramáticas, pare-

ce estar restrito a casos de clivagem, como

(78) a. Fui eu quem comi o peixinho.

b. Fui eu quem comeu o peixinho.

A versão (a) me parece ligeiramente marginal; de qualquer maneira, seria desejável realizar uma investigação mais cuidadosa desses casos.

O segundo caso particular é o de SNs compostos por coordenação. Quando a coordenação se faz com a partícula *e*, ou por justaposição, aplica-se a regra geral que marca o SN como plural:

(74) Antônio e Bia são casados.

(79) Antônio, Bia, Marília, Vera chegaram de trem.

Mas quando a coordenação se faz com certos outros elementos, parece haver vacilação quanto aos traços a serem atribuídos ao SN. Os casos principais são os de *ou* e *nem*. Com *nem*, parece ser preferível aplicar a regra geral:

(80) a. Nem Antônio nem Bia chegaram de trem.

b. ? Nem Antônio nem Bia chegou de trem.

No entanto, a inaceitabilidade de (80b) é muito menor do que, por exemplo, a de

(81) * Antônio e Bia chegou de trem.

Com *ou*, encontra-se tanto o plural quanto o singular. As gramáticas costumam atribuir uma diferença semântica entre as duas versões: no singular, haveria uma idéia de exclusão, e no plural de inclusão:

(82) Antônio ou Bia receberá este cheque.

(83) Antônio ou Bia receberão este cheque.

Em (82) a interpretação seria de que só um deles poderia receber um cheque. Em (83), o cheque poderia ser recebido por Antônio, ou por Bia, ou por ambos. Antes de discutir a possível análise (sintática e semântica) desses exemplos, será necessário apurar se essa oposição de significado é real na língua padrão atual. Por ora, a questão terá de ficar em suspenso.

Finalmente, há o caso freqüentemente citado do "sujeito composto posposto"; em tais casos (sujeito depois do NdP), diz-se que a concordância pode ser feita com o primeiro constituinte do SN composto:

(84) a. Aqui reinam a paz e a alegria.

b. Aqui reina a paz e a alegria.

À primeira vista, parece que teríamos que encontrar um meio de marcar o SN *a paz e a alegria*, opcionalmente, como singular, contrariando assim a regra geral que o marcaria como plural.

No entanto, não me parece evidente que *a paz e a alegria* em (84b) seja realmente um SN. Outra análise, igualmente plausível, pode expli-

car a seqüência como dois SNs um ao lado do outro, não-coordenados; teríamos em (84b) um caso de interpretação anafórica. O sujeito de *reina* seria apenas *a paz*; segue-se um SN que seria sujeito de uma segunda ocorrência de *reina*, caso esta fosse explicitada. Em outras palavras, (84b) seria uma versão anaforicamente reduzida de

(85) Aqui reina a paz e reina a alegria.

Se essa análise for adotada, não haverá problema nenhum a solucionar em (84b); nem existirá o fenômeno da concordância especial com sujeito composto posposto.

No momento, não há evidência decisiva em favor de nenhuma das duas análises. As poucas indicações que conheço favorecem a segunda: primeiro, observe-se que uma vírgula entre os dois SNs em (84b) é muito mais aceitável do que se a seqüência estivesse claramente em função de sujeito:

(86) Aqui reina a paz, e a alegria.

(87) ?? A paz, e a alegria, reinam aqui.

A segunda razão que posso oferecer em favor da segunda análise é que esta nos permite uma descrição mais simples, pois nos livra da necessidade de definir aqui um caso particular de marcação de SNs.

Admitindo embora que esses argumentos não são decisivos, vou adotar a segunda análise, sustentando, portanto, que *a paz e a alegria* em (84b) não formam um SN composto, mas são apenas uma seqüência de dois SNs simples.

6.6.2. Concordância verbal

6.6.2.1. Que é concordância verbal

Nesta seção examinaremos em particular o fenômeno denominado **concordância verbal**. Tradicionalmente, entende-se a concordância verbal como um sistema de condições de harmonização entre o sujeito e o núcleo do predicado das orações. Por exemplo, na oração

(88) Minhas sobrinhas ganharam um cavalo.

essas condições verificariam se o NdP está na forma adequada ao seu sujeito. Já que *minhas sobrinhas* (como vimos em 6.6.1.) é marcado como "terceira pessoa, plural", o NdP precisaria ser preenchido por um verbo que mostrasse as flexões igualmente chamadas de "terceira pessoa, plural". E, como *ganharam* está na terceira pessoa do plural, a frase seria bem formada do ponto de vista da concordância.

Isso, dito assim, fica muito simples, mas na verdade apresenta uma série de problemas tanto teóricos quanto de aplicação aos dados. Nesta seção tentarei discutir os mais

importantes desses problemas, procurando construir um panorama geral do fenômeno da concordância verbal em português. Como se verá, esse panorama difere profundamente da concepção tradicional de concordância verbal; em compensação, é mais coerente e permite uma descrição mais adequada dos fatos da língua.

Nos parágrafos acima, tive o cuidado de qualificar de "tradicional" a concepção da concordância verbal como sistema de harmonização entre o sujeito e o NdP; isso porque, como o leitor verá a seguir, a análise que ofereço desse fenômeno é muito diferente. Para resumi-la em algumas palavras, sustento que não existe propriamente o fenômeno da violação da concordância verbal; isso significa que a inaceitabilidade da frase

(89) * Minhas sobrinhas ganhei um cavalo.

deverá ser explicada por outros meios. E, com efeito, defenderei a idéia de que (89) é excluída por razões outras que não diretamente a desarmonia de pessoa e número entre *minhas sobrinhas* e *ganhei*. A seguir vou apresentar um mecanismo que dá conta desse e de muitos outros fatos usualmente reunidos sob o rótulo de "concordância verbal"; esse mecanismo, como se verá, também explica alguns fatos que a análise tradicional não consegue cobrir.

6.6.2.2. O mecanismo da concordância verbal

A concordância verbal, tal como entendida nesta análise, compreende basicamente alguns procedimentos que rotulam certos sintagmas, atribuindo-lhes funções sintáticas. Esses procedimentos repetem em parte as definições das funções sintáticas estudadas no capítulo 3; mas como não as contradizem, podemos aceitar essa redundância, para efeitos da presente descrição.

Para explicar os casos tradicionalmente considerados como de "violações (ou erros) de concordância", teremos de utilizar, além dos procedimentos de rotulação, alguns outros recursos, que funcionam como filtros. Esses filtros não pertencem propriamente ao mecanismo da concordância, mas são necessários para explicar por que frases como (89) não são aceitáveis. Eles serão expostos na seção 6.6.2.3.

A rotulação se aplica aos SNs de nível oracional (isto é, que não sejam parte de outro sintagma dentro da oração) e que tenham o traço [+Q]. Isso significa, na prática, que se aplicam a sujeitos e objetos diretos. Os procedimentos se formulam da seguinte maneira:

Rotulação dos SNs de nível oracional

(a) Um SN que estiver em relação de concordância com o núcleo do predicado é **sujeito**;

(b) um SN (marcado [+Q]) que não é sujeito é **objeto direto (OD)**;

(c) quanto ao objeto direto:

(c.a) se é um clítico, é um **OD clítico**;

(c.b) se não é clítico, então

(c.b.1.) se vem depois do NdP, é um **OD não-topicalizado**;

(c.b.2.) se vem antes do NdP e contém um elemento Q, é um **OD–Q**; e, se vem antes do NdP, e não contém um elemento Q, é um **OD topicalizado**.

Notas

Os clíticos envolvidos nesses procedimentos são: *o* (com suas variantes de número, gênero e posição: *a, os, as, lo, la, no, na* etc.), além de *me, nos, te, vos.*

Os elementos Q envolvidos são *(o) que, quem, o qual* e SNs introduzidos por *que, qual, quanto(s).*

Como se vê, a novidade consiste apenas em se definir vários tipos de objeto direto: topicalizado ou não, clítico ou não, "Q" ou não.

Vou agora ilustrar brevemente a aplicação desses procedimentos. Seja uma frase como

(88) Minhas sobrinhas ganharam um cavalo.

Os SNs a serem rotulados são *minhas sobrinhas* e *um cavalo.* O SN *minhas sobrinhas* está em relação de concordância com o NdP *ganharam*, e é portanto o sujeito. Passamos a *um cavalo:* verificamos que não está em relação de concordância com o NdP, e é portanto um OD. Não é um clítico e vem depois do NdP — logo, trata-se de um OD não-topicalizado. Se a frase fosse

(90) Um cavalo, minhas sobrinhas ganharam.

um cavalo, por vir antes do NdP e não ser clítico nem conter um elemento Q, seria rotulado de OD topicalizado.

Em

(91) Quantos cavalos sua sobrinha ganhou?

sua sobrinha é o sujeito, e *quantos cavalos* um OD-Q. Finalmente, o OD é clítico em

(92) Minha sobrinha o vendeu.

Note-se que esses procedimentos deixam a possibilidade de não haver sujeito. Isso é necessário para que se obtenha a análise correta em frases como

(93) Ganhei um cavalo.

(94) Está cheio de moscas na cozinha.

Esse é o mecanismo da concordância verbal: como se viu, limita-se a atribuir funções aos SNs de nível oracional.

6.6.2.3. Explicando os "erros de concordância"

O mecanismo visto na seção precedente, como é claro, não conse-

gue explicar a inaceitabilidade de certas frases usualmente consideradas exemplos de "erros de concordância". Acontece que o "erro de concordância" não é uma decorrência direta do mecanismo da concordância, mas de outros fatores gramaticais — em outras palavras, o "erro de concordância" em si não existe. Trata-se, antes, da violação de certos filtros e restrições independentes do mecanismo da concordância. Em um primeiro momento, qualquer combinação (quaisquer SNs acompanhando qualquer forma de verbo no NdP) é bem formada. No entanto, em um segundo estágio, aplicam-se filtros e restrições que marcam como mal formadas muitas dessas combinações; o resultado final, que é o que se observa no uso da língua, é certas combinações serem inaceitáveis.

Examinaremos agora os filtros e restrições mencionados. Podemos começar com o exemplo

(89) * Minhas sobrinhas ganhei um cavalo.

O que é que está errado com essa frase?

Vamos começar rotulando os SNs. Aplicando os procedimentos (a) – (c), descobriremos que *minhas sobrinhas* é um OD topicalizado, e *um cavalo* um OD não-topicalizado. Como se vê, não há sujeito; mas isso não é motivo para rejeitar a frase, porque sabemos que frases sem sujeito são correntes em português.

Entretanto, não é possível construir orações com mais de um objeto direto — o OD é uma das funções não-iteráveis da oração. E (89) tem dois objetos diretos, coisa que nenhum verbo da língua admite (ver as exigências e aceitações dos verbos na seção 6.2.2.). O resultado é que a frase é mal formada — desobedece a transitividade do verbo *ganhar* e também não poderia receber uma interpretação semântica correta, pois não há regra em português que possa atribuir significado a um segundo OD. A conclusão é a seguinte: a má formação de (89) não é um problema de concordância mal feita, mas uma conseqüência de estruturação defeituosa da oração (dois objetos diretos).

Assim, a primeira restrição a ser considerada é a seguinte:

Restrição de transitividade (RT)

A estrutura de uma oração precisa respeitar as condições de transitividade do verbo que ocupa seu NdP.

É fácil verificar que RT é uma condição independentemente necessária. Se podemos estabelecer a transitividade de cada verbo, é justamente porque essa transitividade precisa ser obedecida por todas as orações que contêm esse verbo em seu NdP. (89) é apenas um exemplo específico de aplicação dessa exigência. Conclui-se que a introdução de RT na gramática não se faz com o objetivo específico de dar conta de casos de "erro de concordância";

qualquer gramática do português precisa incluir RT, ou algo equivalente. Como se verá, os demais casos de "erro de concordância" também são explicados com o auxílio de dispositivos independentemente necessários na gramática.

Passemos a um novo exemplo:

(95) * Nós adormeci na banheira.

Rotulando o SN *nós*, verificamos que se trata de um OD topicalizado. O SN *a banheira* não será rotulado, porque não é de nível oracional: é parte de um sintagma maior, *na banheira*, que não é um SN. A oração fica sem sujeito, pois não há nenhum SN em relação de concordância com o NdP.

Ora, se *nós* é um OD, temos aí não uma, mas duas violações. Primeiro, o verbo *adormecer* não aceita OD; e, depois, ainda que aceitasse, a forma *nós* não poderia ser OD, por se tratar de uma forma reta (nominativa). Como resultado, (95) não pode ser aceitável.

Se o verbo fosse um dos que aceitam OD, isto é, se fosse marcado [L-OD] ou [Ex-OD], ainda assim a oração não seria bem formada, porque basta haver uma forma nominativa em função de OD para que a estrutura seja rejeitada. Um exemplo seria:

(96) * Nós encontraram no cinema.

Encontrar é marcado [Ex-OD], e *nós* seria esse OD, preenchendo devidamente as exigências. Mas *nós* não pode ser OD, daí a inaceitabilidade de (96).

Nossa nova restrição, portanto, deverá ser:

Restrição de caso (RC)

Os pronomes pessoais têm formas especializadas quando desempenham a função de OD; nenhuma outra forma desses pronomes pode desempenhar essa função.

Essa restrição está um tanto simplificada aqui, mas serve para nossos objetivos imediatos. É claro que não foi inventada para explicar (95) e (96) — RC é parte de uma condição gramatical do português, e seria necessária ainda que frases como essas fossem aceitáveis.

Um problema diferente é apresentado pela frase

(97) * José mataram os frangos.

Os procedimentos (a – c) darão *os frangos* como sujeito, e *José* como OD topicalizado. Ora, sabemos que *matar* admite objeto direto; por que, então, a frase é mal formada?

Para encontrar a explicação, teremos de fazer uma breve digressão, examinando o seguinte fenômeno: os verbos do português diferem quanto a aceitarem ou não um sujeito após o NdP. Assim, podemos dizer

(98) Chegou meu primo favorito.

mas não

(99) * Ronca meu primo favorito.

Em particular, a posposição do sujeito é impedida quando o verbo tem OD, seja este topicalizado ou não:

(100) * Já escreveu alguns livros esse professor.

(101) * Estes livros, escreveu meu professor.

Os fatores que controlam a possibilidade de pospor o sujeito são mal conhecidos; mas é claro que o verbo *matar* é um dos que não a aceitam:

(102) * Matou o frango José.

Portanto, podemos derivar daí a explicação para a má formação de

(97) * José mataram os frangos.

O único sujeito possível é *os frangos* (pois é o único termo em relação de concordância com o NdP); mas *os frangos* aparece na oração depois do NdP, o que não é permitido: primeiro porque o verbo *matar* não aceita posposição; e, depois, porque a frase tem OD explícito.

A mesma explicação pode ser aventada para

(103) * O traficante encontraram os policiais.

Note-se que todas essas frases — (97), (100), (101), (102), (103) — seriam provavelmente aceitáveis na poesia antiga, como casos de hipérbato; e todas elas podem, na fala, ser muito melhoradas através de artifícios de entoação. No entanto, não são aceitáveis na língua escrita moderna.

Os exemplos examinados e casos paralelos cobrem praticamente toda a gama dos possíveis "erros de concordância" tradicionais. Como se viu, essas violações acabam sendo explicadas através de restrições que não têm nada a ver diretamente com a concordância. Essas restrições atuam como filtros porque "não deixam passar" certas frases que contêm algum tipo de má formação.

6.6.2.4. Vantagens da análise proposta

A análise proposta na seção precedente para o conjunto de fenômenos geralmente considerados como de concordância verbal difere muito da análise tradicional. Se tentarmos examinar uma variedade grande de fenômenos, ficará claro que a nova proposta é superior, pois permite descrever todos aqueles fenômenos da maneira mais simples e coerente. Vamos ver alguns exemplos ilustrativos; naturalmente, a questão continua em aberto, e é possível que se venha a mostrar que, afinal de contas, a solução tradicional é a melhor. No momento, porém, não é isso o que parece.

A solução tradicional se baseia na noção da concordância verbal como **regra**: a forma do verbo é modificada para se harmonizar com os traços de número e pessoa do sujeito.

Já a solução proposta aqui entende a concordância como um sistema de filtros (independentemente motivados) que suprime certas estruturas por apresentarem má formação de algum tipo. É importante enfatizar que esses filtros nunca são específicos dos casos de concordância verbal, mas são necessários na gramática de qualquer maneira.

Vimos que frases como

(104) Comprei um cabrito.

são analisadas nesta *Gramática* como não tendo sujeito. A forma do verbo não é problema, porque segundo nossa análise o verbo não precisa adaptar sua forma a ndo os procedimentos (a) – (c) a (104), temos como resultado que não há sujeito, e há um OD não-topicalizado; nenhuma das restrições vistas na seção 6.6.1.3. se aplica, e portanto a frase é bem formada. Posteriormente, as regras semânticas a processam, atribuindo-lhe uma interpretação na qual o **agente** (não o **sujeito**!) é "eu" (veremos na seção 10.2. como funcionam essas regras semânticas).

Agora, se adotamos a solução tradicional, imediatamente se coloca um problema: com que termo é que o verbo *comprei* está concordando?

Como se sabe, a saída é postular um "sujeito oculto" — a frase (104) conteria o sujeito *eu*, que não se realiza explicitamente, mas que precisa ser sintaticamente real, pois influi na construção formal da oração. Na verdade, esse sujeito só é postulado para possibilitar a manutenção da hipótese da concordância verbal como regra; ele é, portanto, supérfluo, além de introduzir uma abstração na análise.

Às vezes defende-se o sujeito oculto de (104) dizendo-se que ele seria o responsável pela interpretação de um "eu" como agente de *comprei*. Mas o sujeito oculto não é necessário para isso, porque a informação acerca do agente "eu" já é fornecida pela desinência verbal [ver o Quadro 10-C, onde argumento mais detalhadamente contra a existência de sujeito em frases como (104)].

Bem examinadas as coisas, a nova solução é preferível, porque é mais simples, dispensando totalmente a regra de harmonização do verbo com o sujeito, assim como a postulação de um "sujeito" abstrato cuja única função real é sustentar aquela regra de harmonização. Além disso (como argumento no Quadro 10-C), a nova análise permite representar o fato de que a expressão do agente é redundante em (105)

(105) Eu comprei um cabrito.

ao passo que em (104) não é redundante. Ou seja, em (105) a informação de que o agente é "eu" é fornecida simultaneamente pelo sujeito *eu* e pela desinência verbal; daí a redundância. Em (104), essa informação provém de uma única fonte, a desinência.

Passemos a um segundo argumento em favor da nova análise. Uma frase como

(106) Choveu muito em novembro.

apresenta um problema para a análise tradicional, porque é preciso justificar a forma do verbo. Aqui não se pode postular um "sujeito oculto", porque a interpretação semântica esperada ("ele choveu") não se verifica; o sujeito, aliás, não pode ser explicitado:

(107) * Ele choveu muito em novembro.

A análise tradicional nos obriga a criar uma regra especial que coloque o verbo na terceira pessoa do singular sempre que ele não tenha sujeito (claro ou oculto). Naturalmente, a análise adotada aqui não terá problema em explicar por que (106) pode ocorrer. Os procedimentos (a) – (c) não encontrarão sujeito — nem, no caso, OD. E nenhuma das restrições se aplicará à frase, que será bem formada.

Resta explicar a inaceitabilidade de (107). Tradicionalmente, será necessário estipular que *chover* não pode ter sujeito — ou seja, teremos de criar uma restrição especial a esse verbo, somente para explicar por que não ocorrem frases como (107).

Acredito que a nova solução pode dar conta desse fato de maneira mais simples. A análise semântica de *chover* ainda não foi explicitada, mas não há dúvida de que há de incluir a informação de que esse verbo não pode ter um agente; e que pode referir-se a alguma coisa que "cai como chuva", mas que essa coisa teria de ser suficientemente numerosa para imitar uma chuva verdadeira. Evidentemente, a palavra *ele*, que designa uma pessoa ou coisa individual, não se qualifica para ser o elemento que "chove". Mas podemos dizer algo como

(108) Choveram flechas em cima dos pobres coitados.

Em outras palavras, prefiro explicar a inaceitabilidade de (107) como resultado de um traço semântico de *ele*, que resulta em que a interpretação de (107) é mal formada. Não há razões para se afirmar que *chover* não pode ter sujeito [o exemplo (108) mostra que pode]. E de que maneira poderíamos explicar que (108) é aceitável, mas (109) não?

(109) * Choveu uma flecha em cima dos pobres coitados.

É evidente que o problema com (109) não é a presença de um sujeito, mas o significado desse sujeito: uma flecha não basta para dar a impressão — mesmo metafórica — de chuva.

A insistência em vincular a noção de "agente" à presença de um sujeito leva a análises desnecessariamente complicadas e abstratas. Em (104), como vimos, há um "sujeito oculto", no caso vinculado ao agente "eu"; em (106), não há sujeito. Mas

formalmente a situação é a mesma nas duas frases: não há sujeito explícito. As diferenças podem ser explicadas de maneira muito mais natural a partir da semântica dos verbos. Já podemos explicar, por exemplo, por que não se aceita

(110) * Chovi em cima dos meus inimigos.

É porque a forma *chovi* remete a um agente "eu" — singular, individual e, portanto, não adequado a comparações com as gotas de chuva.

Além do "sujeito oculto" de (104) e da ausência de sujeito de (106), a análise tradicional ainda necessita de um terceiro elemento invisível, o "sujeito indeterminado" de

(111) Bateram na porta.

Nessa frase, é preciso admitir que se entende um **agente**, e que este é indeterminado. Mas isso não nos obriga a postular um **sujeito** indeterminado. É perfeitamente possível derivar a interpretação indeterminada (e, acrescente-se, animada e humana) do agente de *bateram* a partir da própria forma verbal; veremos na seção 10.2.6. como se faz isso. Aqui basta observar que um verbo na terceira pessoa do plural, sem sujeito, será interpretado como tendo agente indeterminado.

Dei acima algumas razões que fazem a análise da concordância verbal aqui exposta superior a qualquer análise que tome a concordância como uma regra que altere a forma verbal para fazê-la harmonizar-se aos traços do sujeito. A questão, como já apontei, continua em aberto, pois o fenômeno é bastante complexo; vale a pena examinar novos dados e investigar todas as decorrências da análise, para verificar se esta realmente se sustenta. Por ora, será a análise adotada.

6.6.3. Concordância nominal

6.6.3.1. Que é concordância nominal

Além da concordância verbal, que acabamos de ver, há ainda a **concordância nominal**, assim chamada porque estabelece uma relação morfológica entre elementos tradicionalmente chamados "nomes". A concordância nominal tem certas semelhanças com a concordância verbal, que acabamos de estudar; mas há também diferenças importantes. Podem-se distinguir dois tipos: a concordância entre termos do SN e a concordância de um termo oracional com o sujeito ou o objeto direto; esses dois tipos serão tratados, respectivamente, nas seções 6.6.3.2. e 6.6.3.3.

Em ambos os casos, o fenômeno tem a ver com os traços de gênero e de número de certos constituintes, que precisariam harmonizar-se com os traços correspondentes de

um constituinte considerado central. Assim, temos

(112) Essa mesa nova

onde todos os constituintes estão no feminino singular.

Os constituintes que "concordam nominalmente" são os compostos de um item léxico que **varia em gênero e/ou número**, como *novo* (*nova, novos, novas*) ou *todo* (*toda*). A esses itens atribuiremos os traços [+Ge, +Nu]; trata-se, como se vê, de uma classe de palavras, definida por um critério morfológico (variar em gênero e número) e também sintático (ser levado em conta pelo mecanismo da concordância nominal).

Conforme já vimos, **variar** em gênero precisa distinguir-se de **possuir** gênero: a palavra *novo* **varia** em gênero porque tem uma forma feminina, *nova,* cuja ocorrência é determinada pela sintaxe. Já a palavra *gato* não varia em gênero; *gata* deve ser considerada uma nova palavra (na verdade, um novo **lexema**; ver o capítulo 14), e não uma simples variação de *gato.* Note-se que a ocorrência de *gato* e *gata* não é determinada pela sintaxe, mas depende de uma escolha feita pelo emissor, escolha esta que é determinada pela semântica da mensagem a transmitir.

Isso se reflete no fato de que nem todas as palavras da classe de *gato* têm formas do gênero oposto: *onça, mesa, mão,* por exemplo, não têm masculino. Já as palavras da classe de *novo* sempre têm formas femininas (ainda que às vezes sejam idênticas às masculinas, como em *verde*). Por isso, diremos que a palavra *nova* **está no feminino**, mas a palavra *gata* **é feminina**, ou seja, possui seu próprio item léxico, onde é marcada como feminina.

Itens marcados [–Ge, –Nu] podem ocorrer como modificadores, mas nesse caso não concordam com o núcleo do SN; é o caso de palavras tradicionalmente classificadas como "substantivos", como *menino* em *o imperador menino,* ou de palavras que a gramática tradicional deixa meio em suspenso quanto à classificação, como *rosa, cinza* nos sintagmas *paredes rosa, meu carro cinza.* Observe-se a falta de harmonização quanto a gênero e número entre o modificador e o núcleo do SN.

6.6.3.2. Concordância dentro do SN

A concordância nominal dentro do SN se entende tradicionalmente como um processo que adequa a flexão de certos componentes do SN a traços do núcleo; por exemplo, em

(112) Essa mesa nova

assim como em

(113) Esses livros novos

mas não em

(114) * Essa livro novas

Diz-se, então, que os traços de gênero e número do NSN (*mesa, livros*) fazem com que outros termos do SN, como *esse, novo,* assumam determinadas formas que "concordam" com esses traços.

Essa análise implica em conferir ao núcleo do SN o *status* especial de **cabeça** do SN; os demais elementos precisariam adaptar sua flexão de gênero e número para harmonizar-se com a cabeça do sintagma. Por várias razões, vamos rejeitar aqui essa análise, para adotar uma em que, paralelamente ao que se fez para a concordância verbal, os diversos elementos do SN são associados livremente; em um segundo momento, um sistema de filtros eliminará algumas das seqüências assim criadas, de maneira que (112) e (113) serão produzidos, mas um sintagma como (114) será bloqueado. Uma diferença importante entre a concordância verbal e a nominal é que nesta última o sistema de filtros deverá ser criado especialmente para esse propósito. Desse modo, podemos dizer que a concordância nominal, ao contrário da verbal, é um sistema autônomo; conseqüentemente, o "erro de concordância" existe, quando se trata de concordância nominal, pois é decorrência da violação de um sistema específico de filtros.

Pode-se resumir a concordância nominal que opera dentro dos SNs como um filtro que estipula o seguinte:

Filtro de Concordância no SN
(FCSN)

Marcar como mal formado um SN se houver discordância de gênero e/ou número entre seus constituintes imediatos.

Dois aspectos desse filtro merecem comentário. Primeiro, ele exige que **não haja discordância**, e não que **haja concordância** — isso porque há constituintes do SN que não são capazes, em virtude de sua composição morfológica ou sintática, de mostrar flexão de gênero ou de número. Por exemplo, no sintagma

(115) Cada livro de literatura brasileira

o constituinte *cada* não tem marca de gênero nem de número, por ser invariável (isto é, marcado [–Ge, –Nu]). Igualmente, não é possível marcar gênero ou número no constituinte *de literatura brasileira,* que não constitui um SN. Não se pode, pois, dizer que todos os constituintes de (114) **concordam** em gênero e número; mas pode-se pelo menos afirmar que **não discordam**. Já em

(114) * Essa livro novas

há discordância, porque, por exemplo, *novas* é marcado como feminino e plural, enquanto *essa* é feminino e singular, e *livro* é masculino e singular.

Outro aspecto do filtro que precisa ser explicado é referir-se ele aos constituintes **imediatos** do SN. Isso é para impedir que se marque

como mal formado um sintagma como

(115) Cada livro de literatura brasileira

por ser *brasileira* feminino, e *livro* masculino. Não há discordância aí porque *brasileira* não é um constituinte imediato do SN *cada livro de literatura brasileira*; em vez disso, *brasileira* faz parte do sintagma *de literatura brasileira*, que, ele sim, é um constituinte imediato do SN maior. O filtro de concordância no SN só se aplica aos constituintes imediatamente subordinados ao SN em questão.

Quanto à relevância dos traços [Ge, Nu], ela se evidencia não apenas com palavras do tipo de *cada* [exemplo (115)], mas também com certas palavras que ocorrem no SN e têm gênero e número, mas não variam (são marcadas [–Ge, –Nu]). Um exemplo é

(116) Uma concentração monstro

Note-se que, embora *monstro* seja masculino e *concentração, uma* femininos, não se configura a discordância, porque a palavra *monstro* é [–Ge, –Nu] e, portanto, não é levada em conta pelo filtro.

6.6.3.3. Concordância nominal na oração

Certos termos da oração podem ser ocupados por itens que têm os traços [+Ge, +Nu], isto é, que são passíveis de variação em gênero e número; nesses casos pode também haver concordância com o sujeito ou o objeto direto. Exemplos são as frases

(117) Cristina é rica.

(118) Cristina me procurou furiosa.

(119) Cristina considera Vicente rico.

Em (117) e (118), há concordância de *rica* e *furiosa* com o sujeito *Cristina*; em (119), *rico* concorda com o objeto direto *Vicente*.

O que esses exemplos têm em comum com os casos estudados na seção anterior é a relação morfológica (identidade de traços de gênero e número) entre os constituintes interessados. Mas o fenômeno é bastante diferente sob outros pontos de vista, de modo que mais parece um terceiro tipo de concordância. Aqui continuaremos a chamá-la "nominal", mas acrescentaremos "na oração" para distingui-la da concordância nominal "no SN", que vimos acima.

A concordância nominal na oração também pode ser descrita através de um filtro (distinto do FCSN). Esse filtro se aplica aos constituintes de nível oracional (isto é, constituintes imediatos da oração) que sejam marcados [+Ge, +Nu] e que ocupem funções marcadas [+CN] — ou seja, que sejam complementos do predicado (CP), atributos (Atr) ou predicativos (Pv). A formulação completa do filtro é a seguinte:

Filtro de concordância na oração (FCO)

Marcar como mal formada a oração que tiver um constituinte imediato "C" marcado [+Ge, +Nu], em uma função marcada [+CN], se houver discordância de gênero ou número entre "C" e

(a) o objeto direto (se "C" for predicativo);

(b) o sujeito (se "C" for atributo ou complemento do predicado).

Alguns exemplos devem deixar clara a operação do FCO, que, apesar de seu aspecto meio rebarbativo, é na verdade bastante simples. Seja a frase

(117) Cristina é rica.

Analisando a frase, verificaremos que *Cristina* é o sujeito, e *rica* é complemento do predicado. Sabemos que *rica* é marcado [+Ge, +Nu], pois varia em gênero e número, e que o CP é uma função [+CN]. Portanto, passamos à verificação da existência ou não de discordância.

Como não há predicativo (mas apenas CP), cairemos no caso (b) do filtro, que exige que não haja discordância entre o CP e o sujeito; como é o que acontece (ambos estão no feminino, plural), não há violação, e a frase é bem formada.

Naturalmente, se tivéssemos

(120) * Cristina é rico.

a condição (b) do filtro marcaria a estrutura como mal formada, pois há discordância entre os termos especificados.

Vejamos agora

(121) Eles consideram Cristina rica.

Temos aí um OD, *Cristina*, e um Pv, *rica*; teremos, pois, que aplicar o caso (a) do FCO, que exige que não haja discordância entre o Pv e o OD. Como não há, a frase é bem formada. Já em

(122) * Eles consideram Cristina ricos.

a discordância entre o Pv e o OD causa a má formação. Note-se que *ricos* em (122) não discorda do sujeito *eles*, mas isso não salva a estrutura, porque *ricos* é predicativo, e sua relação de concordância é com o objeto direto.

Um caso aparentemente problemático é o de

(123) Elza encontrou o disco espatifado.

(124) Elza encontrou o disco furiosa.

Em (123), temos um caso paralelo ao de (121), com concordância entre o Pv e o OD. Mas em (124) não teríamos uma frase bem formada onde o Pv (*furiosa*) e o OD (*o disco*) discordam?

A resposta é que *furiosa* é Pv em (123), mas atributo em (124). Isso pode ser verificado pela muito maior facilidade de antepor *furiosa* em (124) (relembro que uma diferença entre o Pv e o Atr é que só este último é marcado [+ Ant]):

(125) ?* Espatifado, Elza encontrou o disco.

(126) Furiosa, Elza encontrou o disco.

Assim, deveremos manter a condição (a), que especifica que o Pv não pode discordar do OD; o sujeito não conta, nesses casos.

Outro exemplo com atributo é a frase

(127) Elza atendeu os clientes gentilíssima.

Se o atributo discordar do sujeito, teremos má formação:

(128) * Elza atendeu os clientes gentilíssimos.

A sentença (128) só será aceitável se entendermos *os clientes gentilíssimos* como um SN; nesse caso, *gentilíssimos* não seria um atributo, evidentemente, mas um modificador, termo interno do SN. Nesse caso, não é possível antepor *gentilíssimos*, nem separá-lo por vírgula. Igualmente, seria difícil colocar um nome próprio no lugar de *os clientes* em (128), porque os nomes próprios, em certas circunstâncias, não aceitam modificadores:

(129) ?? Elza atendeu Ricardo gentilíssimo.

Essa dificuldade não ocorre com (127), onde a substituição **dá** resultado plenamente aceitável:

(130) Elza atendeu Ricardo gentilíssima.

Há casos, a bem dizer bastante duvidosos, de frases com predicativo e sem objeto direto, que colocariam em dificuldade a análise proposta. Um exemplo pode ser a frase

(131) Geralda nasceu magrinha.

Magrinha nessa frase parece um predicativo, e teria os traços [–CV, –Ant, +Q, +CN, +Cl]; no entanto, a relação de concordância se dá com o sujeito, por não haver objeto direto.

Devo observar que a análise de *magrinha* e termos semelhantes como predicativo é algo duvidosa, por causa das flutuações de aceitação do traço [–Ant] nesses casos. O que é claro é que a exigência de não-discordância se aplica aí entre o termo representado por *magrinha* (seja qual for sua função) e o sujeito. Este caso terá de ficar em suspenso até que se possa examinar melhor a função desses constituintes na oração.

6.6.4. Infinitivo flexionado

6.6.4.1. O problema

Um caso especial de concordância verbal é representado pela construção com infinitivo flexionado, exemplificada em

(132) Acho muito triste vocês estarem desempregados.

Essa construção é própria do português, não tendo similar nas outras línguas românicas, e constitui um enigma até hoje não satisfatoriamente analisado, no que pe-

sem os trabalhos já realizados na área (Moraes, 1971; Perini, 1974; Raposo, 1975).

Não parece que o sistema que governa a concordância verbal (isto é, o mecanismo que exclui os casos de discordância) possa dar conta, sozinho, dos casos de infinitivo flexionado; será necessário propor alguns dispositivos específicos para analisar esses casos. Em outras palavras, o fenômeno da concordância do infinitivo não é simplesmente mais um exemplo do fenômeno geral da concordância; o infinitivo segue regras próprias. Nesta seção, proponho uma análise que prevê as ocorrências do infinitivo flexionado (em oposição ao não-flexionado). Aqui, mais do que em qualquer outra seção desta *Gramática*, a solução é meramente descritiva: pretende fornecer um retrato da situação tal como se apresenta na superfície, não chegar a uma explicação integrada do processo dentro do sistema geral da sintaxe da língua. Para isso, deveremos esperar até que se atinja uma compreensão melhor do fenômeno.

6.6.4.2. Papel do verbo principal

Em primeiro lugar, é necessário notar que as possibilidades de ocorrência do infinitivo flexionado dependem de traços (semânticos e sintáticos) do verbo principal, isto é, do verbo que ocupa o NdP da oração que imediatamente contém a oração infinitiva. Assim, com *lamentar* podemos ter

(133) Eles lamentam muitíssimo estarem desempregados.

Mas com **querer** o infinitivo flexionado é impossível:

(134) * Eles queriam muitíssimo vencerem o jogo.

Vamos descrever esse fato através de traços atribuídos aos diferentes verbos. Vou primeiro expor o sistema, e depois darei uma série de exemplos, mostrando como ele funciona para prever as possibilidades de ocorrência do infinitivo flexionado.

Do ponto de vista que nos interessa, os verbos se dividem em três categorias, designadas como tipo A, tipo B e tipo C. O verbo *lamentar* é do tipo A, *querer* é do tipo B, e *ver* é do tipo C. Veremos adiante em que se baseia essa distinção.

6.6.4.3. Prevendo a ocorrência do infinitivo flexionado

Estabelecido isso, podemos passar à formulação das regras que regem o aparecimento de flexão no infinitivo. São apenas três essas regras, complementadas por um filtro, a saber:

1ª Regra do Infinitivo Flexionado (RIF)

Quando o infinitivo é subordinado a um verbo do tipo A, é livre a ocorrência de sufixo de pessoa-número (SPN).

2ª RIF

Quando o infinitivo é subordinado a um verbo do tipo B, a ocorrência de SPN só é admitida se a oração subordinada é introduzida por preposição.

3ª RIF

Quando o infinitivo é subordinado a um verbo do tipo C, a ocorrência de SPN só é admitida se:

(a) a oração subordinada é introduzida por preposição

ou se

(b) há um candidato possível a sujeito da subordinada.

Antes de passar ao filtro, que tem importância secundária, vou exemplificar o funcionamento desse sistema. Tomemos primeiramente um exemplo com verbo do tipo A; com esses verbos, conforme estipula a 1ª RIF, a ocorrência de infinitivo flexionado ou não-flexionado é livre:

(135) a. Os diretores lamentam profundamente não disporem de recursos.
b. Os diretores lamentam profundamente não dispor de recursos.

Nesses casos, ambas as frases são aceitáveis. Para algumas pessoas, (135a) parecerá ligeiramente mal formada; isso se deve ao efeito (variável) do filtro que estudaremos logo adiante.

Vejamos agora um exemplo com verbo do tipo B, como *querer*:

(136) a. * Nós queremos sairmos.
b. Nós queremos sair.

Aqui não há dúvida de que a flexão do infinitivo causa inaceitabilidade. Como o verbo principal é do tipo B, devemos reportar-nos à 2ª RIF; verificamos que a subordinada em (136a), *sairmos*, não é introduzida por preposição; portanto, a presença do sufixo de pessoa-número causa bloqueio da construção. Já (136b), que não tem sufixo de pessoa-número, é aceitável.

Ainda com o verbo *querer* (tipo B), podemos ter

(137) a. Nós queremos um visto para entrarmos no país.
b. Nós queremos um visto para entrar no país.

(137a) é muito mais aceitável que (136a); isso se deve à presença da preposição *para*, que impede que a 2ª RIF exclua a frase.

Vejamos agora um exemplo com verbo do tipo C, como *ver*. Se a subordinada é introduzida por preposição, já sabemos que a flexão é possível:

(138) a. Eles viram o desastre sem fazerem nada.
b. Eles viram o desastre sem fazer nada.

Agora vejamos um caso em que a subordinada vem sem preposição:

(139) a. Vi os cavalos correrem.

b. Vi os cavalos correr.

Ao contrário do que se dá com *querer*, nestas frases a flexão do infinitivo pode ocorrer livremente.

A explicação não é muito fácil; aqui vou dar uma paráfrase da que propus em um trabalho anterior (Perini, 1974). Digamos que, em (139a), seja possível rotular *os cavalos* como sujeito de *correrem*; isto é, como o início da oração subordinada não está formalmente marcado, pode considerar-se como ocorrendo logo depois de *vi*, assim:

[Vi [os cavalos correrem]]

Nesse caso, o objeto direto de *vi* será oracional. A oração subordinada será então sujeita às condições normais de concordância verbal, isto é, *os cavalos* será rotulado como sujeito; e, como o verbo *correr* não exige nenhum outro complemento, a estrutura é bem formada.

Já em (139b) a mesma segmentação, isto é,

[Vi [os cavalos correr]]

resultará em má formação, pois *os cavalos* será rotulado como objeto direto, e esse sintagma não é um objeto adequado ao verbo *correr*. No entanto, admitimos que também está disponível a segmentação

[Vi os cavalos [correr]]

que, como se pode facilmente verificar, dá resultado bem formado (esse fenômeno é chamado em Perini, 1974, **flutuação**).

Se aceitarmos, portanto, a possibilidade de dupla análise para essas seqüências (flutuação), teremos uma explicação para a aceitabilidade de (139a) e (139b): em (139a) há um candidato possível a sujeito da subordinada, o que faz com que a estrutura escape à exclusão.

Note-se que o mesmo não pode ocorrer com verbos do tipo B, que não admitem flutuação:

(140) * Nós queremos nós sairmos.

Com estes verbos, a presença de um candidato a sujeito da subordinada (o segundo *nós*) não é relevante, pois sua aceitabilidade é governada pela 2ª RIF; por isso (140) é inaceitável.

Voltando a considerar os exemplos de (139), observa-se que se, por alguma modificação na estrutura, deixa de haver um candidato a sujeito da subordinada, torna-se impossível a ocorrência do infinitivo flexionado. É o que acontece se no lugar de *os cavalos* colocamos um clítico:

(141) * Vi-os correrem.

Ao contrário de *os cavalos*, *-os* não poderia ser sujeito (pois as formas oblíquas são especializadas em funções não-subjetivas). Por isso, a flexão na subordinada fica impedida. Natural-

mente, o infinitivo sem flexão continua aceitável:

(142) Vi-os correr.

Para os falantes que aceitam a passiva, na oração principal de frases do tipo de (139) dá-se o mesmo:

(143) * Os cavalos foram vistos correrem.

Como *os cavalos* terá de ser rotulado como sujeito da principal, não fica disponível para ser sujeito da subordinada; conseqüentemente, a 3ª RIF marca a frase como mal formada. Novamente, o infinitivo sem flexão dá resultado aceitável (para os falantes que aceitam a passiva na principal):

(144) Os cavalos foram vistos correr.

Passemos agora a examinar um problema que surge com frases como

(145) ? Eles lamentam morarem tão longe.

(146) ? Nós lamentamos morarmos tão longe.

Como *lamentar* é um verbo do tipo A, a 1ª RIF estabelece que a flexão do infinitivo é livre, e (145) e (146) deveriam ser aceitáveis; no entanto, percebe-se alguma estranheza nelas, sem que se possa dizer que são tão mal formadas quanto, por exemplo, (136a). Essa estranheza decresce quando se introduz material entre o verbo principal e o subordinado:

(147) Eles lamentam profundamente morarem tão longe.

(148) Nós lamentamos do fundo do coração morarmos tão longe.

Por outro lado, a estranheza só surge quando o verbo principal e o infinitivo têm flexão de pessoa-número idêntica:

(149) Eles lamentam morarmos tão longe.

Vou descrever esses fatos através de um filtro, a saber:

Filtro de Dupla Desinência
(FDD)

Marcar como semi-aceitáveis as frases que contenham um infinitivo flexionado precedido de uma forma verbal com SPN idêntico ao do infinitivo.

O FDD resulta em inaceitabilidade relativamente leve, motivo pelo qual optei por colocar "semi-aceitáveis" na formulação. Além disso, o efeito do FDD decresce com o aumento da distância entre os dois verbos envolvidos.

A ação do FDD explica por que frases como

(135) a. Os diretores lamentam profundamente não disporem de recursos.

podem parecer ligeiramente marginais para alguns falantes. Naturalmente, é de esperar que a frase piore se retiramos a palavra *profundamente*.

6.6.4.4. Exemplificação

As três regras vistas acima, mais o Filtro de Dupla Desinência, constituem o mecanismo que trata especificamente da ocorrência do infinitivo

flexionado em português. Conforme já fiz notar, permanecem problemas sem solução, alguns deles advindos da própria incerteza dos falantes quanto à aceitabilidade de muitas frases com infinitivo flexionado; mas acredito que o mecanismo proposto dá conta da imensa maioria dos casos. A seguir forneço mais uns poucos exemplos de sua aplicação.

Comecemos examinando as frases

(150) a. Eu trouxe o frango para vocês o assarem.
b. * Eu trouxe o frango para vocês o assar.

Aqui a situação parece o oposto da habitual, pois o infinitivo só pode ocorrer flexionado; vejamos por quê.

(150a) não apresenta problema: como a oração de infinitivo é introduzida por preposição, a ocorrência do sufixo de pessoa-número é livre, qualquer que seja o tipo a que pertence o verbo. Mas por que (150b) será inaceitável?

As RIFs, assim como o Filtro de Dupla Desinência, não podem excluir (150b). Mas a oração subordinada de (150b), como qualquer oração da língua, está ainda sujeita às outras condições relevantes. Assim, temos que aplicar à oração *vocês o assar* os procedimentos de rotulação vistos na seção 6.6.2. O resultado será que o pronome *o* é objeto direto e que *vocês* é igualmente objeto direto (pois não está em relação de concordância com o verbo). Essa situação, como vimos na seção mencionada, é intolerável, pois nenhuma oração da língua pode ter mais de um objeto direto. Como resultado, *vocês o assar* é mal formada — o que acarreta, claro, a má formação de (150b). Em (150a), a rotulação é diferente, porque *vocês* está em relação de concordância com *assarem* e, portanto, é rotulado como sujeito; só há um objeto direto, e a construção é bem formada.

No exemplo

(151) a. * Os portugueses foram à África procurarem ouro.
b. Os portugueses foram à África procurar ouro.

temos um infinitivo sem preposição e sem candidato possível a sujeito; conseqüentemente, a 2ª RIF marca (151a) como mal formada, pois aí o infinitivo está flexionado; (151b), com infinitivo sem flexão, é bem formada. Ao que tudo indica, pois, *ir* deve ser um verbo do tipo B ou C.

Vejamos agora

(152) ? Eles prometeram a Bira chegarem na hora.

Se admitirmos que *prometer* é um verbo do tipo A, a flexão do infinitivo será livre; o resultado é uma frase aceitável. Por outro lado, como o verbo principal e o infinitivo são identicamente flexionados para pessoa e número, o FDD marcará a frase como semi-aceitável.

O verbo *prometer* é algo excepcional de mais de um ponto de vista, e apresenta alguma dificuldade também quanto à ocorrência do infinitivo flexionado em seu complemento. A frase seguinte

(153) ? Eles prometeram (a Bira) chegarmos na hora.

é bastante estranha, embora segundo a análise proposta devesse ser perfeitamente aceitável — já que é idêntica a (152), exceto que o SPN é diferente nos dois verbos e, portanto, não sofre a ação do FDD. Não conheço no momento explicação satisfatória para esse fato.

6.6.4.5. Os três tipos de verbos: algumas notas

Resta fazer algumas considerações sobre a divisão dos verbos em três grupos, segundo seu comportamento em face da ocorrência de um infinitivo flexionado em seu complemento. Essa subclassificação, naturalmente, tem algo de arbitrário, pois não se conhecem outros traços sintáticos importantes que dividam o conjunto dos verbos nesses mesmos três grupos. Por outro lado, é necessário considerar o caráter bastante marginal do infinitivo flexionado dentro da sintaxe do português: é um tempo verbal cujas condições de concordância diferem das dos outros tempos (que, estes, se comportam todos da mesma maneira); talvez por isso mesmo, seu uso se vem reduzindo, e as intuições dos falantes a respeito das condições de sua ocorrência são muito menos seguras do que para os outros tempos verbais.

Não obstante, algumas observações podem ser feitas a respeito dos três tipos. Comecemos pelo tipo A, que inclui os verbos *lamentar, revelar, ignorar, perceber, compreender, prometer.* É interessante notar que, no tipo A, estão praticamente todos os verbos chamados **factivos**, isto é, aqueles que apresentam a propriedade semântica de pressupor a verdade de seu complemento. Quando se utiliza um verbo factivo, fica entendido (entre o falante e o ouvinte) que o significado da oração subordinada é verdadeiro; comparem-se

(154) Elza ignorava que André a traía.

(155) Elza dizia que André a traía.

Em (154) fica entendido que a traição de André é um fato, ao passo que em (155) ela é apenas a opinião expressa de Elza.

Como se vê, temos aqui uma correlação aproximada entre uma categoria semântica (verbos factivos) e uma categoria sintática (verbos do tipo A). A correlação é aproximada porque, de um lado, pelo menos um verbo factivo, *saber*, não admite livremente a flexão do infinitivo, parecendo pertencer ao tipo B; e, de outro lado, há alguns poucos verbos não-factivos que perten-

cem ao tipo A; vimos que é o caso de *prometer*.

Vejamos agora o tipo B, exemplificado por *querer*. Esses são os verbos que têm subjuntivo no complemento sempre que não se entende identidade entre os elementos apontados pelos sufixos de pessoa-número (e/ou sujeitos) dos dois verbos:

(156) O chefe quer que (nós) fiquemos até as oito.

Caso haja identidade entre esses elementos, o subjuntivo é inaceitável:

(157) * Nós queremos que (nós) fiquemos até as oito.

Esses verbos admitem complemento com infinitivo (sem flexão), formando frases que de certa forma preenchem a lacuna deixada pela impossibilidade de correferência mencionada:

(158) Nós queremos ficar até as oito.

O tipo C, finalmente, compreende os verbos que não se encaixam em nenhum dos tipos acima; provavelmente, trata-se do caso mais comum na língua.

7 Sistemas de correspondência

7.1. A RELAÇÃO DE CORRESPONDÊNCIA

7.1.1. Definições

Na seção 2.1.3.5., apresentou-se a noção de **correspondência**, que, como se viu, serve para descrever certas relações formais entre estruturas sintáticas diferentes. Por exemplo, são **correspondentes** as estruturas com topicalização e suas versões não-topicalizadas, como

(1) a. Leo comeu o peixinho.
b. O peixinho, Leo comeu.

Esses relacionamentos são relevantes tanto para a descrição morfossintática quanto para a interpretação semântica. São importantes para a descrição morfossintática porque expressam implicações de aceitabilidade entre estruturas formalmente distintas; ou seja, a partir de certas estruturas aceitáveis, pode-se prever a existência na língua de outras estruturas aceitáveis.

Isso, entre outras coisas, permite uma simplificação notável na enumeração das estruturas da língua. Por exemplo, se estabelecermos que a toda frase com objeto direto corresponde uma outra frase com o objeto direto topicalizado, só precisaremos enumerar as estruturas da língua que contêm objeto direto não-topicalizado [como, por exemplo, (1a)]; as versões topicalizadas [como (1b)] serão automaticamente derivadas das primeiras. Além disso, como veremos, o elemento topicalizado pode ser considerado como tendo a mesma função [isto é, *o peixinho* em (1b) é objeto direto, tal como em (1a)], o que igualmente contribui para simplificar a análise.

Por outro lado, a relação de correspondência tem importância para a interpretação semântica porque é possível discernir uma sistematicidade na relação de significado entre frases correspondentes. Elas são "semanticamente semelhantes", em um sentido que ainda não foi inteiramente explicitado. No caso de (1a) e (1b), pode-se notar que ambas as frases têm o

mesmo conteúdo, embora sejam diferentes quanto ao elemento que é tomado como tópico da mensagem.

A relação de correspondência está presente, de uma forma ou de outra, na maioria dos modelos de análise sintática. Desse modo, ao a adotarmos, não fazemos mais que seguir uma tendência geral (não propriamente universal, pois muitos lingüistas a rejeitam). No Quadro 7-A, faço um breve apanhado da presença dessa relação na literatura gramatical.

Costuma ser fácil detectar a correspondência entre frases, em parte porque esta é acompanhada por semelhança semântica; a lista de exemplos que veremos adiante deverá ajudar o leitor nessa tarefa. Mas é necessário definir formalmente a relação de correspondência, para que seja possível decidir em casos limítrofes e para que fique clara a distinção entre "correspondência" e, de outro lado, "semelhança semântica" ou "paráfrase". Passo, portanto, à definição formal dessa relação.

A correspondência pode ser **total** (A corresponde a B, e vice-versa) ou **parcial** (A corresponde a B, mas B não corresponde a A). Ambas as modalidades deverão ser definidas aqui; começarei pela correspondência total.

Correspondência total
(definição)

Duas formas A e B serão totalmente correspondentes se for possível estabelecer entre todos *os termos de A e* todos *os termos de B um relacionamento um-a-um tal que:*

(i) os membros de cada par assim formado sejam preenchidos por itens léxicos idênticos; e

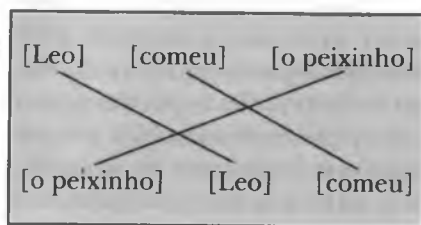
(ii) para qualquer preenchimento léxico idêntico dos pares, a aceitabilidade de A implique a aceitabilidade de B, e vice-versa; e a inaceitabilidade de A implique a inaceitabilidade de B, e vice-versa.

Adendo

Quando se diz "todos os termos de A e todos os termos de B", admite-se a exceção dos elementos ser ... que *e* o que ... ser, *típicos das construções clivadas.*

Vejamos alguns exemplos da aplicação dessa definição. Em primeiro lugar, sejam os exemplos

(1) a. Leo comeu o peixinho.
b. O peixinho, Leo comeu.

A primeira coisa a fazer é estabelecer uma relação um-a-um entre os termos das duas frases. Para isso, naturalmente, selecionaremos os termos identicamente preenchidos, conforme indicado no esquema

Agora observe-se o seguinte: qualquer substituição paralela de itens dará resultado idêntico nas

duas frases, quanto à aceitabilidade. Assim, se substituirmos, nas duas, *Leo* por *Carvalho*, ambas as frases continuarão aceitáveis. Mas, se substituirmos *Leo* por *dormíamos*, ou por *a gramática portuguesa*, ambas as frases ficarão inaceitáveis. O importante é o seguinte: **nenhuma substituição paralela pode dar resultado diferente nas duas frases**. Concluímos que (1a) e (1b) são totalmente correspondentes.

Examinemos agora as frases

(2) Maria adora Geraldo.

(3) Geraldo adora Maria.

Essas duas frases são, à primeira vista, tão estreitamente relacionadas entre elas quanto (1a) e (1b). No entanto, como veremos, não são correspondentes.

Estabelecidas as relações entre os termos identicamente preenchidos, passamos ao teste das substituições paralelas. É claro que, se substituirmos *Maria* por *Magda*, ambas as frases continuarão aceitáveis. Mas se substituirmos *Maria* por *essa sonata de Mozart*, o resultado será diferente:

(4) * Essa sonata de Mozart adora Geraldo.

(5) Geraldo adora essa sonata de Mozart.

Ou seja, não se verifica aqui a implicação de aceitabilidade exigida pela definição de correspondência — o que significa que (2) e (3) não são frases correspondentes.

Como se vê, a aplicação da definição não é tão complexa como pode parecer à primeira vista. Passemos agora à definição de "correspondência parcial".

Correspondência parcial
(definição)

Uma forma A será parcialmente correspondente a outra forma B se for possível estabelecer entre todos *os termos de B e* alguns *termos de A um relacionamento um-a-um tal que:*

(i) os membros de cada par assim formado sejam preenchidos por itens léxicos idênticos; e

(ii) para qualquer preenchimento léxico idêntico dos pares, a aceitabilidade de A implique a aceitabilidade de B (mas não vice-versa).

O exemplo mais conhecido é o dos pares ativa/passiva, como

(6) a. Geraldo estragou um saxofone.
b. Um saxofone foi estragado por Geraldo.

O que a definição realmente estipula é que para cada passiva existe sempre uma ativa, mas para certas ativas não existem passivas. Falando mais formalmente, qualquer termo que colocarmos no lugar dos termos de (6b) (com resultado aceitável) também deve poder ser colocado no lugar do termo correlato de (6a). Assim, se substituirmos, em (6b), *um saxofone* por *meu fusquinha*, o resultado será aceitável em ambas as frases:

(7) a. Geraldo estragou meu fusquinha.
b. Meu fusquinha foi estragado por Geraldo.

Isso acontece com todos os termos de (6b) que têm correlatos em (6a). Se o resultado for inaceitável em (6b), naturalmente será inaceitável em (6a):

(8) a. * Geraldo dormiu meu fusquinha.
b. * Meu fusquinha foi dormido por Geraldo.

Mas há casos em que a substituição dá certo na ativa, mas não na passiva, como

(9) a. Geraldo tem um saxofone.
b. * Um saxofone é tido por Geraldo.

Por isso dizemos que (6b) corresponde a (6a), mas não vice-versa. Generalizando, diremos que as frases passivas correspondem a ativas, mas as ativas não correspondem a passivas. Ou seja, a correspondência aqui é parcial.

7.1.2. Condicionamento gramatical e condicionamento léxico

A diferença entre correspondência total e correspondência parcial decorre de uma diferença entre dois tipos de relação estrutural: relações dependentes apenas da estrutura e relações dependentes (pelo menos em parte) de propriedades de itens léxicos individuais.

Como exemplo de relação puramente estrutural, tomemos novamente o caso das topicalizações:

(1) a. Leo comeu o peixinho.
b. O peixinho, Leo comeu.

Podemos definir a previsão de aceitabilidade em termos puramente estruturais: a toda oração com a estrutura **Sujeito + Predicado + Objeto direto** corresponde necessariamente outra oração, igualmente aceitável, com a estrutura **Objeto direto + Sujeito + Predicado**. Isso não depende dos itens léxicos que estejam desempenhando as funções mencionadas; quaisquer que sejam eles, a implicação de aceitabilidade (e portanto a relação de correspondência) é válida. A relação se define exclusivamente em termos das estruturas envolvidas; esses são casos de correspondência total.

Já no caso de

(6) a. Geraldo estragou um saxofone.
b. Um saxofone foi estragado por Geraldo.

a implicação de aceitabilidade depende do verbo que esteja preenchendo o núcleo do predicado: se na ativa o verbo for *estragar*, existirá uma passiva correspondente; mas, se o verbo for *ter*, a passiva não será aceitável. A relação se define, portanto, não apenas em termos das estruturas envolvidas, mas também em

termos dos itens léxicos individuais presentes na oração; esses casos são de correspondência parcial.

Podemos traduzir esses exemplos como regras gerais do seguinte modo: observando exemplos como (1a) e (1b), diremos que "um objeto direto pode sempre ser transportado para o início da oração principal" (isto é, pode ser topicalizado). Nenhum item léxico precisa ser mencionado. Exprimindo o mesmo fato de outro modo, podemos afirmar em geral que o objeto direto, independentemente dos itens que o preencham, tem o traço [+Ant].

Observando os exemplos do tipo de (6a) e (6b), teremos de fazer uma afirmação algo diferente: "a cada oração ativa corresponde uma passiva, **se o NdP for preenchido por determinados verbos**: *estragar, comprar, examinar,* mas não *ter*". Como se vê, é essencial mencionar aqui os itens léxicos individuais. Naturalmente, essa informação está contida nos itens léxicos correspondentes: no item léxico de *ter*, encontramos a informação de que esse verbo "não admite passiva".

Exprimimos essa diferença dizendo que a correspondência total é uma relação **condicionada gramaticalmente**, e a correspondência parcial é (em parte) **condicionada lexicalmente**. O condicionamento nos casos de correspondência parcial não é inteiramente léxico, é claro, porque mesmo aí é necessário mencionar as estruturas envolvidas.

7.1.3. Funções sintáticas e a correspondência

Examinando um grupo de orações correspondentes como

(1) a. Leo comeu o peixinho.

b. O peixinho, Leo comeu.

pode-se perguntar como se devem analisar sintaticamente: terão ambas o mesmo conjunto de funções sintáticas, ou as diferenças formais existentes entre elas acarretam diferenças de função?

Vou adiantar aqui a resposta, que terei de comentar adiante: nos casos de correspondência total, as frases têm sempre as mesmas funções sintáticas, ocupadas pelos mesmos elementos léxicos — exceção feita apenas dos elementos mencionados no adendo da definição, isto é, as marcas de clivagem. Estas, como só existem em uma das versões, naturalmente não poderiam ser analisadas da mesma maneira nas duas. Já nos casos de correspondência parcial pode haver mudanças de função de certos termos. Nos parágrafos seguintes, vou apenas dar a análise adotada nesta *Gramática*; as razões para adotar essa solução, e não outra, estão no Quadro 7-B.

Vejamos primeiramente as razões para a manutenção das funções nos casos de correspondência total. Voltando ao exemplo (1), isso quer dizer que *o peixinho*, objeto direto em (1a), deve ser também analisado

como objeto direto em (1b). Ora, isso é decorrência automática da própria definição de "objeto direto". Conforme vimos no capítulo 3, o OD é definido por certos traços, entre os quais [+Ant] — isto é, o OD pode ser anteposto (topicalizado). No exemplo (1b), temos exatamente um OD anteposto, o que é previsto pela definição. É claro que, se *o peixinho* em (1b) não fosse OD, teríamos de modificar a definição dessa função (assim como os procedimentos de rotulação utilizados para identificá-la como "OD topicalizado" — ver a seção 6.6.2.).

Em outras palavras, a identidade de funções de *o peixinho* nessas duas orações está incluída na definição de OD. Quando se diz que o OD é [+Ant], isso significa que ele pode sempre ocorrer anteposto **em uma oração correspondente**. Esse princípio vale para todos os casos de correspondência total; assim, *o peixinho* é objeto direto em

(10) Foi o peixinho que Leo comeu.

porque o OD é também marcado [+Cl].

Já nos casos de correspondência parcial pode haver diferenças de função. Nos pares passiva/ativa, como

(11) a. Os mecânicos estragaram meu fusquinha.
b. Meu fusquinha foi estragado pelos mecânicos.

a própria definição das funções nos obriga, como é óbvio, a definir *os mecânicos* como sujeito de (11a), mas *meu fusquinha* como sujeito de (11b). Desse modo, o princípio de identidade de funções não é válido para casos de correspondência parcial.

7.2. GRUPOS DE CORRESPONDÊNCIA TOTAL

7.2.1. Definições

Chamaremos **grupo de correspondência** um conjunto de estruturas que se correspondem total ou parcialmente.

A relação de correspondência, como se pode ver pelas definições dadas em 7.1.1., se aplica a estruturas individuais, e não a estruturas generalizadas; por exemplo, as definições permitem verificar que as frases

(1) a. Leo comeu o peixinho.
b. O peixinho, Leo comeu.

são totalmente correspondentes; mas as definições não informam nada diretamente sobre a correspondência entre estruturas das formas

(12) a. Suj + NdP + OD
b. OD, + Suj + NdP

Verifica-se, no entanto, que toda oração individual da forma (12a) é totalmente correspondente a uma

oração individual da forma (12b). Como não há exceções, podemos fazer a afirmação mais geral de que as estruturas [Suj + NdP + OD] e [OD + Suj + NdP], elas próprias, são totalmente correspondentes. São afirmações generalizadas como essa que devem ser incluídas na gramática; os casos listados abaixo serão considerados sob esse ponto de vista, e os grupos de correspondência serão definidos em termos dos constituintes e das funções sintáticas das orações envolvidas.

Devo alertar o leitor para o fato de que a lista de grupos de correspondência (em especial no que diz respeito às correspondências parciais) é provisória, pois houve pouca pesquisa sistemática nesse campo. A correspondência, como já apontei acima, embora seja uma noção presente na maioria das análises, não é em geral explicitada, e certamente nunca se elaborou uma lista completa dos grupos de correspondência para o português. Uma lista preliminar, para o inglês, encontra-se no artigo básico de Harris, 1964.

Veremos a seguir uma lista (provavelmente incompleta) dos grupos de correspondência existentes em português. Em alguns casos, tentei definir com exatidão a relação que existe entre os membros do grupo; em outros casos, preferi não tentar uma descrição rigorosa, por sentir que seria prematuro fazê-lo.

7.2.2. Topicalização

A formulação da topicalização que dei acima não está completa, porque a relação funciona mesmo se houver outros termos além dos três mencionados; por exemplo, o caso de

(13) a. Leo comeu o peixinho imediatamente.

b. O peixinho, Leo comeu imediatamente.

deve ser considerado como sintaticamente paralelo ao de (1a)/(1b). Além disso, também se pode colocar no mesmo caso grupos onde se topicaliza outro termo que não o objeto direto:

(14) a. Leo está irritado.

b. Irritado, Leo está.

Aqui se topicalizou o complemento do predicado, mas a relação pode ser considerada a mesma. Na verdade, a relação exemplificada pelos exemplos (1), (12) e (13) vale sempre que houver na oração um termo marcado [+ Ant]. Os outros termos da oração não são relevantes, a não ser para marcar a posição relativa do elemento topicalizado. Podemos então definir a relação de topicalização assim:

Topicalização

Sempre que em uma oração houver um termo marcado [+Ant], haverá também uma oração idêntica a essa,

com a diferença de que o termo em questão está no início da frase.

7.2.3. Anteposição de elemento Q

Vimos, na seção 6.6.2.2., que um **elemento Q** é um dos itens de uma pequena lista: *que, o que, quem, quando, como, onde, por que.* Esses elementos desempenham as funções sintáticas usuais na oração e têm a propriedade de ocorrerem em duas posições: ou no início do período, ou então na posição que seria de esperar que ocupassem, dada sua função. Por exemplo, temos pares de sentenças como

(15) a. Vocês procuram o quê?
b. O que vocês procuram?

ou então

(16) a. O Marcelo mora onde?
b. Onde o Marcelo mora?

(Ocorrem outras modificações de ordem, mas não temos que ocuparnos delas aqui; por exemplo, também se diz *onde mora o Marcelo?*)

Os pares de (15) e de (16) constituem grupos de correspondência, pois em cada caso as duas frases se correspondem totalmente. Devemos portanto formular para eles uma relação de correspondência, a saber,

Anteposição de elemento Q

Sempre que em uma oração houver um elemento Q, haverá também uma oração idêntica a essa, com a diferença de que o elemento Q está no início da frase.

O leitor provavelmente terá notado que essa relação é muito semelhante à topicalização; alguns autores (como Decat, 1978) já sugeriram que se trataria do mesmo fenômeno.

Uma nota a respeito da topicalização e da anteposição de elemento Q: a anteposição do constituinte se faz levando-o para o início da **frase** (período), e não para o início da oração. Isso se evidencia em casos em que o elemento pertence a uma oração subordinada não-inicial:

(17) a. O peixinho, Leo disse que Mabel comeu.
b. ? Leo disse que o peixinho, Mabel comeu.

(18) a. Onde vocês disseram que Mabel morava?
b. * Vocês disseram que onde Mabel morava?

Conforme indicado acima, parece que a violação resultante de antepor o constituinte para o início da oração (subordinada) é mais séria no caso da anteposição de elemento Q do que no caso da topicalização; de qualquer modo, a anteposição para o início do período é sempre preferível.

7.2.4. Clivagem

As frases clivadas já são nossas conhecidas, assim como o fato de que correspondem totalmente a suas versões não-clivadas. Um exemplo é o par

(19) a. Mamãe fez os sanduíches.
b. Foi mamãe que/quem fez os sanduíches.

A relação formal é relativamente complexa. Envolve o transporte do elemento clivado, o acréscimo do verbo *ser* no mesmo tempo do verbo original e o acréscimo de *que* ou *quem.* A seleção destes dois últimos elementos depende de traços do constituinte clivado; se se tratar de sintagma marcado [+humano], pode-se usar *que* ou *quem,* indiferentemente, como está em (19b). Mas, se o constituinte clivado for marcado [–humano], só se pode usar *que*:

(20) a. Foi o bolo que me fez mal.
b. * Foi o bolo quem me fez mal.

(O traço [humano] é provavelmente de caráter semântico; ou seja, é provável que a má formação de (20b) derive de alguma incompatibilidade entre a semântica de *quem* e a de *bolo.*)

A clivagem tem uma versão, menos freqüente, em que o constituinte transportado se coloca no início da construção:

(21) Mamãe foi que/quem fez os sanduíches.

Uma possível interpretação dessa construção seria a de que se trata de uma aplicação à mesma estrutura da topicalização e da clivagem. Que essa aplicação dupla é possível se percebe por exemplos como

(22) Os sanduíches, foi mamãe que/quem fez.

Nesta frase, topicalizou-se o SN *os sanduíches* e clivou-se o SN *mamãe.* Em (21), *mamãe* teria sido clivado e também topicalizado. Levando em conta essa análise, não incluirei a possibilidade de (21) na formulação da clivagem. Tampouco incluirei a seleção de *que* ou *quem,* admitindo que se trate de fenômeno semântico.

A clivagem pode, portanto, ser descrita pela definição seguinte:

Clivagem

Sempre que em uma oração houver um termo marcado [+Cl], haverá também uma outra oração que difere da primeira nos seguintes particulares:

(a) a oração começa com o verbo ser, *no mesmo tempo em que está o verbo principal da oração primitiva;*

(b) segue-se o termo marcado [+Cl];

(c) segue-se o item que(m)*;*

(d) seguem-se, na ordem, os demais termos da oração primitiva.

Apesar da evidente complexidade da relação, as orações clivadas são de uso muito corrente tanto na fala quanto na escrita.

7.2.5. Pseudoclivagem

É o nome que se dá a uma relação que tem semelhanças com a clivagem, sem se identificar com ela. Por exemplo, sejam as sentenças

(23) a. Vovô assou o cabrito.

b. Quem assou o cabrito foi vovô.

À primeira vista, parece ser apenas uma variante da clivagem, com mudança de ordem dos constituintes. Mas além dessa há outras diferenças bastante curiosas. Primeiro, não existe a opção quanto ao elemento Q inicial, que na clivagem pode ser *que* ou *quem*: nestes casos, ele é totalmente determinado pelo traço [humano] do constituinte que é colocado em evidência. Assim, em (23b), não se pode colocar *que* no lugar de *quem*:

(24) * Que/o que assou o cabrito foi vovô.

Outra diferença entre as duas construções está em que, enquanto a clivagem se aplica a constituintes marcados [+Cl], a pseudoclivagem só se aplica a constituintes marcados [+Q]. Assim, se tomarmos, por exemplo, um adjunto circunstancial, que é [+Cl, –Q], poderemos clivá-lo, mas a versão pseudoclivada não será aceitável:

(25) Maria foi entrevistada por Sílvio.

(26) Foi por Sílvio que Maria foi entrevistada.

(27) * Quem/o que/por quem Maria foi entrevistada foi (por) Sílvio.

O mesmo se repete para todas as funções marcadas [+Q]: a pseudoclivagem só é possível com o sujeito, o objeto direto, o predicativo e o complemento do predicado, que são as funções marcadas [+Q]. Por conseguinte, temos de formular a relação de pseudoclivagem separadamente da de clivagem, a saber:

Pseudoclivagem

Sempre que houver em uma oração um termo marcado [+Q], haverá também outra oração que difere da primeira nos seguintes particulares:

(a) começa com que *ou* quem*;*

(b) segue-se o restante da oração, sem o constituinte marcado [+Q];

(c) segue-se o verbo ser, *no mesmo tempo do verbo principal da oração primitiva;*

(d) segue-se o constituinte marcado [+Q].

7.2.6. Movimentação de clíticos

Outro caso de correspondência total é o de frases que só diferem quanto à posição de um clítico ("pronome oblíquo", na nomenclatura tradicional):

(28) a. Suas idéias me assustam.

b. Suas idéias assustam-me.

(29) a. Mariquinha se estava maquiando.
b. Mariquinha estava-se maquiando.
c. Mariquinha estava se maquiando.
d. Mariquinha estava maquiando-se.

As regras que governam o posicionamento dos clíticos serão vistas no próximo capítulo (seção 8.1.4.). É preciso observar que a liberdade de movimentação dos clíticos é algo que tende a desaparecer da língua. Em cada um dos grupos acima, apenas uma versão é plenamente aceitável no padrão atual: (28a) e (29c); as outras são, em graus variáveis, todas um pouco "corretas demais", ou seja, arcaizantes.

De qualquer modo, dentro de um registro que aceite todas essas frases, não há dúvida de que (28) e (29) são grupos de correspondência total: não há, por exemplo, verbos ou pronomes que exigem ênclise ou próclise; as condições são puramente sintáticas, nunca léxicas.

7.2.7. Movimentação do predeterminante

Vimos na seção 4.3. que o predeterminante é um elemento externo ao SN propriamente dito e que pode ocorrer em diversas posições na oração. As possibilidades de movimentação do PDet foram especificadas naquela seção. Por exemplo, temos:

(30) a. Todas as adolescentes gostam de tango.
b. As adolescentes todas gostam de tango.
c. As adolescentes gostam todas de tango.

Como o leitor poderá verificar, trata-se de um caso bem claro de correspondência total.

7.2.8. Movimentação de Atr e AO

Vimos no capítulo 3 que o adjunto oracional e o atributo se caracterizam por marcas positivas quanto a traços de movimentação dentro da oração: são marcados [+Ant, +PA] (outros traços, naturalmente, distinguem o AO do Atr). Esses traços definem automaticamente grupos de correspondência total. Podemos considerar que o traço [+Ant] define grupos de topicalização; assim, o caso de

(31) a. Nosso diretor viaja freqüentemente.
b. Freqüentemente, nosso diretor viaja.

se encaixaria na topicalização, vista acima em 7.2.2. Mas, para o traço [+PA], será necessário definir um novo tipo de correspondência total:

c. Nosso diretor freqüentemente viaja.

(31c) se vincula a (31a) e (31b) através de uma relação de correspondência a que chamarei **movimentação de Atr e AO**.

7.3. GRUPOS DE CORRESPONDÊNCIA PARCIAL

Vejamos agora os grupos de correspondência parcial mais importantes. Relembro o leitor que, nos casos de correspondência parcial, é possível prever a existência na língua de uma frase ou sintagma B a partir da existência de outra frase ou sintagma A, mas não vice-versa. Assim, dada uma oração passiva bem formada, sabe-se que "sua" ativa é bem formada; mas, dada uma ativa, pode acontecer que "sua" passiva não seja bem formada.

7.3.1. Ativas e passivas

O caso paradigmático de correspondência parcial é o dos pares ativa/passiva. Para cada passiva, há sempre uma oração ativa correspondente; mas, como vimos, nem toda ativa tem uma passiva correspondente. Isso se deve, geralmente (mas não sempre), a traços do verbo. Assim, os verbos *cozinhar, fazer, escrever* e *ler* entram tanto em construções passivas quanto em ativas. Já *ter* só pode entrar em construções ativas (se excluirmos seu emprego especial no sentido de "considerar": *ele é tido por louco*). Não há verbos que só possam ocorrer em construções passivas. Para efeito da descrição, vamos admitir que os verbos que admitem passiva são marcados com o traço [+Pass], e os que não admitem são marcados com o traço [–Pass].

A relação passiva/ativa se define da seguinte forma:

Passivização

Sempre que houver uma oração formada por

> *um sujeito* SN_1
>
> *+ uma forma do verbo* ser
>
> *+ o particípio de um verbo* V_1, *marcado [+Pass]*
>
> *+ um adjunto circunstancial formado da preposição* por + SN_2,

haverá também uma oração formada de

> SN_2 *como sujeito*
>
> *+* V_1 *no tempo e modo de* ser *na primeira estrutura*
>
> *+* SN_1 *como objeto direto.*

[A primeira estrutura se denomina **passiva**; a segunda, **ativa**.]

Como se vê, a relação é das mais complexas, o que não impede o uso intensivo e freqüente de orações passivas. Para exemplificar, tomemos a frase

(32) Tomás foi surpreendido pelos detetives.

Essa oração satisfaz a descrição da passiva, dada acima: é formada de um sujeito (*Tomás*), seguido de uma forma de *ser* (*foi*), seguido do particípio de um verbo [+Pass] (*surpreendido*), seguido de um AC com a preposição *por* (*pelos detetives*). Assim, prevemos que deve existir na língua uma oração aceitável formada de *os detetives* como sujeito, o verbo *surpreender* no mesmo tempo e modo (não necessariamente número e pessoa!) de *foi*, seguido de *Tomás* como objeto direto. A previsão é confirmada, pois (33) é aceitável:

(33) Os detetives surpreenderam Tomás.

7.3.2. Alçamento de objeto

Pares de frases como

(34) Mastigar farinha é difícil.

(35) Farinha é difícil de mastigar.

também formam grupos de correspondência parcial: (35) é parcialmente correspondente de (34). Essa construção é governada por um traço do adjetivo que ocupa a função de complemento do predicado: no caso, *difícil*. Se mudarmos o adjetivo em (34), poderá não haver frase correspondente:

(36) Mastigar farinha é humilhante.

(37) ?? Farinha é humilhante de mastigar.

Vou chamar essa relação de **alçamento de objeto**, mantendo a metáfora transformacional de que o objeto direto de (34) é "alçado" à posição de sujeito de (35). Adjetivos como *difícil* são marcados com o traço [+AlçO].

A relação se define assim:

Alçamento de objeto

Sempre que houver uma oração com a estrutura seguinte

> *sujeito formado de oração reduzida de infinitivo,* O_1*, sem sujeito e com objeto direto*
>
> *+ o verbo* ser, estar *ou* parecer
>
> *+ um CP representado por um adjetivo marcado [+ AlçO],*

haverá também uma estrutura formada de

> *o objeto direto de* O_1*, como sujeito*
>
> *+ o NdP e o CP da primeira estrutura*
>
> *+ a preposição* de
>
> *+ o verbo de* O_1.

Novamente uma relação bastante complexa — e novamente um tipo bastante freqüente de oração. O leitor poderá verificar por si mesmo que as descrições se aplicam a exemplos como (34) e (35).

Ocorrem, no português coloquial, certas construções que se assemelham a (35), sem serem idênticas a ela. Provavelmente representam relações de correspon-

dência próximas do alçamento de objeto; por exemplo, muitas pessoas aceitam um tipo de alçamento em que O_1 tem sujeito (e pode dispensar o objeto) na estrutura não-alçada:

(38) Ernesto é difícil de sair à noite.

que corresponderia a

(39) Ernesto sair à noite é difícil.

Não vou incluir esse caso aqui porque (38) me parece nitidamente coloquial, e nesta *Gramática* estamos estudando apenas o português padrão (o que não implica, obviamente, em negar interesse ao estudo dessas e das outras construções do português coloquial).

Outra possível versão do alçamento é representada pelo par

(40) Ernesto ler este livro é difícil.

(41) Este livro é difícil para Ernesto ler.

Não estou certo de que se trate de correspondência; é um caso a examinar.

7.3.3. Pré-núcleos e modificadores

Muitos itens léxicos (assim como sintagmas) podem desempenhar mais de uma função; mas isso nem sempre determina uma relação de correspondência. Por exemplo, sabemos que tanto *Zé* quanto *Elisa* podem ser sujeitos ou objetos diretos; no entanto, é fácil mostrar que não há relação de correspondência entre as frases

(42) Zé viu Elisa.

(43) Elisa viu Zé.

Consultar, a respeito, a seção 7.1.1.

No entanto, parece que estruturas que contêm pré-núcleo interno mais NSN são parcialmente correspondentes de estruturas em que o mesmo item ocupa a função de modificador externo; por exemplo, (44) parece ser correspondente de (45):

(44) Um lindo dia

(45) Um dia lindo

Isso se deve ao fato de que todos os itens que podem ser PNI podem também ser ModE. Como muitos itens podem ser ModE mas não PNI, a implicação de aceitabilidade é unidirecional e a correspondência é parcial: (44) corresponde a (45), mas não vice-versa. A relação pode ser definida assim:

Anteposição de adjetivo

Sempre que houver um SN que contenha um item I na função de PNI, haverá também um SN idêntico a esse, mas com I na função de ModE.

7.4. POSPOSIÇÃO DE SUJEITO

Sabemos que, em muitas circunstâncias, o sujeito de uma oração pode ocorrer após o NdP:

(46) Chegaram vários carregamentos de cachaça.

(47) Onde está Laurinha?

(48) É triste que você não tenha sido reeleito.

Os fatores que governam a posposição do sujeito são ainda misteriosos, no que pesem vários trabalhos importantes feitos sobre o assunto. Já foram propostos fatores sintáticos, léxicos, semânticos e discursivos, sem (que eu saiba) se chegar a uma solução cabal. Aqui vou apenas apontar a existência do fenômeno, evitando propor uma análise — no que estou recuando da posição mais arrojada que adotei anteriormente (Perini, 1989, p. 83). É que, pensando melhor, acho que ainda não compreendi adequadamente o fenômeno.

Parece-me claro, todavia, que a posposição define uma relação de correspondência; e certamente há casos em que a posposição é impossível:

(49) * Comeu uma pizza Sônia.

Uma das restrições sintáticas à posposição de sujeito é que esta não pode ocorrer na presença de um objeto direto, como em (49).

Como não se sabe ao certo o papel que desempenham na posposição de sujeito as idiossincrasias léxicas, não é claro se se trata de correspondência parcial ou total. Vamos deixar a questão, como tantas outras, em suspenso, até que seja suficientemente esclarecida.

Há certamente outros grupos que merecem ser examinados, para ver se manifestam outras relações de correspondência; na verdade, mal se começa a compreender o fenômeno, no que pese sua presença nas descrições lingüísticas, há já longo tempo.

7.5. A CORRESPONDÊNCIA NA DESCRIÇÃO DA LÍNGUA

Para concluir, vou sumariar aqui as razões que sustentam o uso das relações de correspondência na descrição gramatical.

Em primeiro lugar, observemos que as implicações de aceitabilidade são um fato da sintaxe da língua e só por isso merecem figurar na sua descrição. É um fato, por exemplo, que, para cada frase composta de **sujeito + NdP + objeto direto**, existe outra frase, igualmente aceitável, composta de **objeto direto + sujeito + NdP**. O que se pode discutir é o grau de utilidade dessa relação para o restante da descrição da língua; aqui, como se viu, estou apostando em que a relação é especialmente útil.

Uma vez definidas as relações de correspondência, torna-se possível formular as transitividades verbais de maneira mais econômica. Assim, não será necessário dizer, ao formularmos a transitividade do verbo *fazer*, que este exige objeto direto em estruturas como (50), mas dispensa o objeto direto em estruturas como (51) ou (52):

(50) Minha firma fez a estrutura desse edifício.

(51) A estrutura desse edifício foi feita por minha firma.

(52) Uma estrutura é difícil de fazer.

Em vez disso, observamos que tanto (51) quanto (52) são frases parcialmente correspondentes: (51) de (50), e (52) de (53)

(53) Fazer uma estrutura é difícil.

E estipulamos que as transitividades se formulam apenas para estruturas **não-parcialmente correspondentes**. A complementação dos verbos nestas é verificada derivativamente, a partir da transitividade do verbo **mais** a definição da relação de correspondência em questão. Isso nos permite manter a afirmação simples de que *fazer* recusa objeto direto; e analogamente permite simplificar a transitividade de muitos outros verbos.

Essa é, em resumo, a motivação para se incluir na gramática as relações de correspondência. Há, a bem dizer, muitos problemas não solucionados acerca dessa relação, e há quem a considere dispensável, ou mesmo insustentável. No momento, porém, a solução aqui adotada me parece a mais adequada aos interesses da descrição da sintaxe da língua. Relembro que as relações de correspondência, ou algo muito próximo delas, estão presentes na maioria dos modelos atuais de sintaxe; esses modelos são, nesse sentido, todos transformacionais em alguma medida. Há autores que não incluem relações análogas à correspondência em seus modelos, mas estes são uma minoria (o que não quer dizer que não possam ter razão! A questão está muito longe de uma decisão final).

8 Estruturas sintáticas do português

8.1. ORDEM DOS TERMOS NA ORAÇÃO

8.1.1. Regras de estrutura sintagmática

A ordem dos termos dentro da oração já está em parte contida nas próprias definições das funções, assim como na explicitação das relações de correspondência. Assim, sabemos que se pode dizer

(1) a. Rogério vai fazer o churrasco.
b. O churrasco, Rogério vai fazer.

Essas duas possibilidades decorrem automaticamente do fato de o objeto direto ser marcado [+Ant], podendo, pois, ser colocado no início da frase.

Em particular, os traços [Ant], [PA] e [pNdP] especificam possibilidades de ordenação (ver as definições no capítulo 3). As relações de correspondência, como vimos no capítulo 7, são também em grande parte uma simples reafirmação desses traços: podemos, por exemplo, descrever (1a) e (1b) como um caso de correspondência total (topicalização).

Por outro lado, o elenco de traços utilizados para basear a análise sintática desenvolvida no capítulo 3 não inclui todos os traços possíveis. Ali só houve a preocupação de distinguir claramente as diversas funções — e, a bem dizer, nem todas as funções importantes foram definidas. Isso significa que muitos outros traços possivelmente importantes não foram explicitados nesta análise; caso os explicitássemos, provavelmente obteríamos, já pronta, boa parte da descrição das ordens de termos possíveis na oração portuguesa.

Partindo dos traços [Ant] e [PA], podemos distinguir dois tipos principais de termos: os que aceitam ser colocados no início da frase (além de em uma outra posição da oração), e que são marcados [+Ant]; e os de posicionamento livre dentro da oração, que são marcados [+PA]. Vamos examinar brevemente cada um desses casos.

O caso mais típico de termo marcado [+Ant] é o do objeto direto. Como se pode ver nos exemplos de (1), o OD pode ocorrer no início da frase (o que decorre de seu traço [+Ant]), ou então logo após o NdP. Conforme já notei acima, esta última possibilidade poderia ser expressa em um novo traço, que seria parte de uma definição mais completa de OD. Essa definição completa encerraria então uma descrição de todas as posições em que o OD pode ocorrer dentro da oração. E, se repetirmos essa complementação das definições para todos os termos da oração, teremos, ao mesmo tempo, uma descrição das possibilidades de ordenação desses termos dentro da oração.

Os termos marcados [+PA] podem ser considerados como de posição livre, no sentido de que ocorrem livremente dentro da oração. Um exemplo é o atributo:

(2) a. Ana Maria vai à Europa todo ano.
b. Ana Maria vai todo ano à Europa.
c. Ana Maria todo ano vai à Europa.
d. Todo ano Ana Maria vai à Europa.

Essas frases se correspondem totalmente, e mostram que o atributo (*todo ano*) tem posição livre dentro da oração. Isso não significa que possa ocorrer **dentro** dos constituintes oracionais, é claro:

(3) * Ana Maria vai à todo ano Europa.

A posição livre se define apenas em termos dos constituintes de nível oracional.

Esses dois traços, além de [pNdP], que exprime a posição fixa da negação verbal, cobrem parte das variações de posição encontradas na oração. Mas existe, além disso, um grande número de outras restrições de ordenação (que poderiam, em princípio, ser descritas também através de traços). Vamos examinar um exemplo simples, baseado na frase (1a).

Essa frase se compõe de **sujeito + predicado + objeto direto**, sendo que o predicado é complexo e se compõe de **auxiliar + núcleo do predicado**. Sabemos que pelo menos duas ordenações desses elementos são possíveis, a saber, [Suj + Aux + NdP + OD] (frase (1a)) e [OD, + Suj + Aux + NdP] (frase (1b)). Muitas outras ordenações são possíveis em princípio, e é necessário apurar quais delas são também aceitáveis.

O resultado será que **nenhuma** é aceitável. Dou abaixo apenas alguns exemplos:

(4) a. * Vai Rogério fazer o churrasco.
b. * Vai fazer Rogério o churrasco.
c. * Vai fazer o churrasco Rogério.
d. * Rogério fazer vai o churrasco.
e. * Rogério fazer o churrasco vai.
f. * O churrasco vai fazer Rogério.

[Esta é inaceitável enquanto variante de (1a); poderia ser aceitável como

estrutura independente, mas semanticamente anômala.]

g. * O churrasco Rogério fazer vai.
h. * Vai Rogério fazer o churrasco.
i. * Vai fazer Rogério o churrasco.
j. * Fazer Rogério vai o churrasco.
l. * Fazer vai o churrasco Rogério.

etc., etc., etc.

É importante observar que as ordenações exemplificadas em (4) são inaceitáveis quaisquer que sejam os itens léxicos envolvidos. Por exemplo, nenhuma substituição de verbos poderia fazer com que alguma dessas frases se tornasse aceitável; trata-se, portanto, de inaceitabilidades gramaticalmente condicionadas, e não lexicalmente condicionadas. Vimos no capítulo 7 que nem sempre é esse o caso; em frases sem objeto direto, alguns verbos admitem a posposição do sujeito, outros não. Selecionei, para o presente exemplo, um caso particularmente simples e claro; em outros casos surgem complicações.

Voltando aos exemplos de (4), podemos observar que a sintaxe impõe restrições extremamente rigorosas à ordenação dos termos da oração. No caso examinado, dentre vinte e quatro ordenações possíveis dos quatro elementos, nada menos que vinte e duas são inaceitáveis.

É evidente que, nesse caso, fica muito mais econômico formular as seqüências possíveis do que as restrições de ocorrência. Diremos, portanto, que em português são possíveis as seguintes combinações desses quatro elementos (quando sozinhos em uma oração):

[Suj + {Aux + NdP} + OD]
[OD, + Suj + {Aux + NdP}]

Podemos representar isso mais compactamente da seguinte maneira: primeiro, observamos que {Aux + NdP} é simplesmente a estrutura interna do predicado. Assim, a estrutura da oração se simplifica como

[Suj + Pred + OD]
[OD, + Suj + Pred]

Separadamente, a estrutura do predicado será:

[Aux + NdP]

ou melhor, já que o Aux não está obrigatoriamente presente,

[(Aux) + NdP]

onde os parênteses indicam que a ocorrência do Aux é opcional.

Em segundo lugar, observamos que a ocorrência da segunda estrutura oracional mencionada, [OD, + Suj + Pred], já é prevista pela relação de correspondência total que denominamos **topicalização** (seção 7.2.2.) — ou, alternativamente, pelo traço [+ Ant] do OD. Assim, basta-nos informar que a estrutura da oração (com sujeito, objeto direto e predicado) é

[Suj + Pred + OD]

nessa ordem.

Uma última observação relevante é que tanto o sujeito quanto o

objeto direto podem estar ausentes de uma oração:

(5) Comprei um filhote de dálmata. (sem sujeito)

(6) Rogério adormeceu. (sem objeto direto)

Assim, colocaremos também esses dois elementos entre parênteses na fórmula que descreve a oração. Dessa maneira, ficamos com as duas fórmulas seguintes:

para a oração:

[(Suj) + Pred + (OD)]

para o predicado:

[(Aux) + NdP].

Essas fórmulas são na verdade **regras**, isto é, instruções sobre o modo de construir as estruturas em questão. Vamos, então, grafá-las da maneira usual em sintaxe, a saber,

RES-1: O → [(Suj) + Pred + (OD)]

RES-2: Pred → [(Aux) + NdP].

Entenda-se: (a) uma **oração** se compõe de um **sujeito** (opcional), seguido de um **predicado** (obrigatório), seguido de um **objeto direto** (opcional). E (b) um **predicado** se compõe de um auxiliar (opcional) seguido de um núcleo do predicado (obrigatório). Essas regras recebem o nome de regras de estrutura sintagmática (e por isso as designo como "RES") e têm como função especificar as estruturas possíveis na língua.

Uma pergunta que poderia ser feita é: Por que RES-1 explicita diretamente a estrutura não-topicalizada (com o OD depois do Pred)? Por que não fazer uma regra que explicite a versão topicalizada (OD + Suj + Pred), derivando daí a forma sem topicalização? A resposta é que, partindo da forma sem topicalização, chegaremos a uma análise geral mais simples; por exemplo, incluindo na regra geral os casos de OD representado por pronomes clíticos, que nunca podem ocorrer topicalizados. Há mais o que dizer a respeito disso, mas aqui não será possível discutir o assunto.

É óbvio que a primeira regra não é suficiente para descrever todas as estruturas oracionais do português; mas é suficiente para descrever todas as orações que compreendem apenas um ou mais dos seguintes termos: sujeito, predicado e objeto direto. Quanto à segunda regra, ela deixa de consignar a possibilidade de ocorrência de mais de um auxiliar, como em

(7) Rogério tem estado trabalhando em seu livro.

Para ilustrar o funcionamento do sistema, porém, vou limitar-me a essas duas regras, formuladas como estão. Veremos como elas contribuem para explicar uma grande variedade de estruturas, bem e mal formadas.

8.1.2. Estrutura da oração: um exemplo

Considerando apenas três termos, o sujeito, o predicado e o objeto direto, a regra de estrutura sintagmática que especifica as orações possíveis em português é apenas

RES-1: O → [(Suj) + Pred + (OD)].

Nenhuma outra regra será necessária. Mas, naturalmente, teremos de levar em conta outros componentes da sintaxe (e da semântica) da língua, tais como as relações de correspondência e vários dos filtros já estudados. Com esse arsenal de dispositivos, poderemos explicar as aceitabilidades e inaceitabilidades observadas.

Em primeiro lugar, observemos que RES-1 permite produzir ("gerar") quatro estruturas. A primeira é

[Suj + Pred + OD]. Exemplo:

(1) a. Rogério vai fazer o churrasco.

A segunda é **[Pred + OD]**:

(5) Comprei um filhote de dálmata.

A terceira é **[Suj + Pred]**:

(6) Rogério adormeceu.

E a quarta é **[Pred]**:

(8) Chegamos.

Retomemos agora a oração

(1) b. O churrasco, Rogério vai fazer.

Essa oração não se encaixa nas previsões de RES-1, que só admite o sujeito em primeiro lugar e o objeto direto depois do predicado. Mas (1b) é prevista pela relação de topicalização, vista na seção 7.2.2. Como o OD é marcado [+Ant], sabemos, pela topicalização, que, se uma estrutura da forma [Suj + Pred + OD] é aceitável, necessariamente também a língua admite [OD, + Suj + Pred]. Portanto, (1b) é prevista pela gramática — através da ação conjunta de RES-1, mais a relação de topicalização.

Isso se repete, obviamente, para todas as relações relevantes vistas no capítulo 7. Assim, a aceitabilidade de

(1) a. Rogério vai fazer o churrasco.

nos permite prever automaticamente a aceitabilidade de (1b) e também de

(9) É o churrasco que Rogério vai fazer. (**clivagem**)

(10) É Rogério que vai fazer o churrasco. (**clivagem**)

(11) Quem vai fazer o churrasco é Rogério. (**pseudoclivagem**)

(12) O que Rogério vai fazer é o churrasco. (**pseudoclivagem**)

(13) O que Rogério vai fazer? (**anteposição de elemento Q**)

(14) Rogério o vai fazer. (**movimento de clítico**)

e assim por diante.

Agora examinemos o caso de

(15) * Rogério vai fazer.

Como, segundo RES-1, o OD é opcional, essa oração, composta apenas de sujeito e predicado, deveria ser bem formada. Mas uma estrutura só é aceitável quando atende a **todas** as exigências relevantes, não apenas às regras de estrutura sintagmática. E acontece que (15), embora não desobedeça RES-1, contém uma violação à Restrição de Transitividade (vista na seção 6.6.2.3.). Com efeito, o verbo *fazer* é marcado como [Ex-OD], isto é, exige objeto direto; e como o predicado de (14) contém o verbo *fazer* e nenhum OD, a estrutura é marcada como mal formada, sendo, conseqüentemente, inaceitável.

Passemos agora ao exemplo

(16) * Eu vai fazer o churrasco.

Essa oração parece ser composta de sujeito, predicado e objeto direto e deveria, portanto, ser bem formada. No entanto, os procedimentos de rotulação vistos no capítulo 6 rotulam tanto o SN *o churrasco* quanto o SN *eu* como objetos diretos (ver a seção 6.6.2.2.). Isso resulta em duas violações: primeiro, (16) tem dois objetos diretos, o que não pode acontecer com nenhuma oração; e, depois, um desses objetos diretos é a forma *eu*, que não pode desempenhar essa função (só *me* pode ser OD) — trata-se de uma violação à Restrição de Caso, formulada na seção 6.6.2.3. Por isso, a oração é mal formada.

Já a seqüência

(17) * Rogério o churrasco vai fazer.

é mal formada por desobedecer RES-1 (a ordem dos termos está incorreta); além disso, não há qualquer relação de correspondência que se aplique a essa seqüência, salvando-a da inaceitabilidade. Como o OD é marcado [-PA], não pode nunca ocorrer entre o sujeito e o predicado, como está em (17).

8.1.3. De onde vem a ordenação

Sumariando o que se viu até o momento, pode-se dizer que a ordem dos termos na oração é produto de um complexo de fatores. Em primeiro lugar, temos regras que, entre outras coisas, estipulam a ordenação dos termos; são as regras de estrutura sintagmática, como RES-1. Em segundo lugar, há condições sintáticas como a que marca como mal formada qualquer oração que contenha mais de um objeto direto; ou a que exclui uma oração onde a transitividade do verbo não é satisfeita.

Uma oração será bem formada (do ponto de vista da ordem de seus termos) quando atender a duas condições: (a) a ordenação dos termos for prevista diretamente por alguma RES ou, então, por alguma RES em conjunção com alguma relação de correspondência; e (b) nenhuma restrição sintática a marcar como mal formada.

8.1.4. Posicionamento dos clíticos

8.1.4.1. O problema

Um caso especial é constituído pelos clíticos (tradicionalmente chamados "pronomes oblíquos"), ou seja, as formas *me, o, se, lhe, nos* (além das pouco usadas *te, vos*). Essas formas são flexões dos itens *eu, ele, nós* (*tu, vós*) e se posicionam dentro da oração segundo princípios próprios.

Os princípios, como veremos, são bastante simples (no que pesem as grandes discussões que às vezes suscitam). O verdadeiro problema está nas freqüentes incertezas de julgamento quanto à posição dos clíticos em certos casos — decorrência do fato de que, nesse ponto, as variedades brasileiras diferem muito do padrão europeu, causando vacilação constante entre a tendência a respeitar esse padrão e a tendência a adaptá-lo ao nosso uso. No que se segue, considerarei um compromisso que me parece ser o mais amplamente aceito pelos usuários do padrão brasileiro.

Distinguem-se duas posições dos clíticos, definidas estas em relação ao NdP ou ao Aux: **próclise** (colocação do clítico imediatamente antes do NdP ou Aux), como, por exemplo, em

(18) O sermão me aborreceu.

(19) O sermão me vai aborrecer.

e **ênclise** (colocação do clítico imediatamente depois do NdP ou Aux), como em

(20) Aborreci-me com o sermão.

É bom notar que os clíticos sempre ocorrem obrigatoriamente contíguos ao verbo (Aux ou NdP), seja antes, seja depois dele. Quando ocorrem depois, ligam-se ao verbo por hífen.

A chamada **mesóclise** é apenas um caso especial de ênclise, que aparece quando o NdP ou o Aux está no futuro do presente ou do pretérito; as condições em que se admite a ênclise valem igualmente para a mesóclise. Um exemplo de mesóclise é

(21) Aborrecer-me-ei com o sermão.

O problema principal é, portanto, o de determinar as circunstâncias em que a língua admite a próclise e/ou a ênclise.

8.1.4.2. Restrições à próclise e à ênclise

As duas restrições seguintes bastam para prever a imensa maioria dos casos:

Posicionamento dos clíticos

Restrição à próclise:

É mal formada toda oração que contenha proclítico no início de estrutura oracional não-subordinada ou logo após elemento topicalizado.

Restrição à ênclise:

É mal formada toda oração que contenha enclítico quando:

o elemento verbal (Aux ou NdP) é gerúndio, precedido de em;

ou

o Aux/NdP é particípio;

ou

a oração se inicie com item marcado [+Atração].

Em todos os outros casos, usa-se próclise ou ênclise, indiferentemente.

Isso é tudo. A única dificuldade consiste em se saber exatamente quais são os itens marcados [+Atração]; isso veremos abaixo.

Antes de exemplificar o funcionamento das duas restrições, chamarei a atenção do leitor para o seguinte fato: a ênclise está desaparecendo do português brasileiro; essa tendência, dominante na modalidade falada, já deixou marcas muito profundas no próprio padrão escrito. Conseqüentemente, as duas restrições tendem a desaparecer, para serem substituídas por um princípio mais simples, que estabelece apenas que os clíticos se colocam sempre antes do NdP. A análise acima cobre uma forma muito conservadora da língua, conforme se verá na exemplificação. É necessário atualizá-la, mas isso deverá ser precedido de um levantamento do uso dos clíticos no padrão brasileiro moderno; por ora, contentemo-nos com as observações informais dadas nesta seção.

Vejamos agora alguns exemplos da aplicação das restrições à próclise e à ênclise. A Restrição à Próclise exclui frases como as seguintes:

(22) * Me preocupei com vocês.

(23) * Me dê uma ajuda aqui.

(24) * Ontem à noite, me comportei mal.

(25) * Telefonei várias vezes e me preocupei com vocês.

O leitor observará que todas essas frases ocorrem normalmente na língua falada e mesmo com freqüência na escrita. De qualquer forma, serão consideradas mal formadas em uma modalidade da língua (muito conservadora) que inclua a Restrição à Próclise tal como formulada acima. Uma modalidade intermediária provavelmente aceitaria (25), rejeitando (22), (23) e (24).

Quando a oração é subordinada, a próclise é possível (ou mesmo obrigatória, por força da Restrição à Ênclise):

(26) Eu já disse que me preocupei com vocês.

A Restrição à Ênclise vai marcar como mal formadas frases como:

(27) * Em aceitando-se sua proposta, devemos adiar a votação.

(28) * Ele tem comportado-se mal.

(29) * Ele não apresentou-se ao comitê.

(Considerando-se que *não* é um item marcado [+Atração].)

Nos outros casos, aceita-se próclise ou ênclise; por exemplo,

(30) a. Minhas primas comportam-se bem.
b. Minhas primas se comportam bem.

Quando o predicado é complexo, pode-se fazer ênclise ao Aux ou ao NdP, desde que não se apliquem as restrições à ênclise:

(31) a. Minhas primas estão-se comportando bem.
b. Minhas primas estão comportando-se bem.

No padrão brasileiro, pode-se igualmente fazer a próclise a qualquer dos dois:

(32) a. Minhas primas se estão comportando bem.
b. Minhas primas estão se comportando bem.

A próclise ao NdP, quando há Aux, era condenada pelos gramáticos antigos; não se sabe até que ponto essa condenação refletia um uso real no padrão escrito da época. Desse modo, (32a) seria considerada "correta", mas (32b) seria "incorreta". No padrão brasileiro atual, ambas são aceitáveis, e a próclise ao NdP [como em (32b)] é pelo menos mais freqüente do que a próclise ao Aux [como em (32a)]. A próclise ao NdP é também mais freqüente do que a ênclise a qualquer dos dois elementos, como em (31a) e (31b), de modo que se pode dizer que a posição natural do clítico, quando o predicado é complexo, é a próclise ao NdP: precisamente a construção antigamente considerada incorreta.

8.1.4.3. *Casos de "atração"*

Vamos considerar agora o fenômeno da "atração", exercido por certos itens léxicos aos quais atribuí o traço [+Atração]. Todas as gramáticas mencionam o fato de que certos itens "atraem" o clítico, de maneira que quando esses itens iniciam a oração somente a próclise é possível; no entanto, não há muito consenso sobre quais são exatamente esses itens. Os relativos e interrogativos são sempre mencionados:

(33) a. O gato que me arranhou era cinza.
b. * O gato que arranhou-me era cinza.

(34) a. Que gato o arranhou?
b. * Que gato arranhou-o?

Outros elementos que constam de todas as listas são:

— o item *não*;
— *nunca, só, até, mesmo, também*;
— *tudo, nada, alguém, ninguém*;
— o complementizador *que*.

Às vezes também são mencionados

— SNs acompanhados de predeterminante (*todos os rapazes, ambos os rapazes*);

— SNs iniciados por *qualquer, nenhum*.

Finalmente, algumas gramáticas citam ainda outros itens; por exemplo, Cunha & Cintra (1985) consignam *bem, mal, ainda, já, sempre.*

A única maneira de decidir quais desses itens realmente exercem atração é fazer levantamentos nos textos, assim como testes com pessoas que dominam o padrão. Por outro lado, não há dúvida de que o fenômeno realmente existe, pois (33b) e (34b), pelo menos, são claramente inaceitáveis. Na falta de estudos detalhados sobre o assunto, teremos de deixar a lista em aberto; fica a sugestão de pesquisa.

8.1.4.4. Até onde vai a próclise?

Vimos que em casos de próclise o clítico se coloca logo antes do NdP:

(35) O gato me arranhou.

E, quando há uma seqüência de auxiliar + NdP, pode-se fazer a próclise ao auxiliar ou então (no padrão brasileiro atual) também ao NdP:

(36) a. O gato me está arranhando.

b. O gato está me arranhando.

A construção exemplificada em (36a) se faz cada vez mais rara no padrão brasileiro. No entanto, ainda se encontra, e é interessante ver quando pode ocorrer.

Em geral, pode-se dizer que o clítico se posiciona antes do auxiliar, ou seja, fica sempre restrito aos limites da mesma oração. Assim, a frase (37a) não tem a variante (37b):

(37) a. O sujeito demonstrou interessar-se pelo tema.

b. * O sujeito se demonstrou interessar pelo tema.

Isso pode ser atribuído ao fato de que o clítico de *interessar-se* está colocado em outra oração.

No entanto, há alguns casos duvidosos, que é necessário pelo menos mencionar aqui. Primeiro, com o verbo *saber* (que não é um auxiliar), encontra-se às vezes o clítico transportado:

(38) Ela me sabe agradar.

Essa construção, embora rara, ainda ocorre por vezes.

Depois, o clítico antes do Aux tem aceitabilidade mais ou menos reduzida quando o Aux vem acompanhado de preposição, como em

(39) ?? Ela se deixou de maquiar.

(40) * Ela se continuou a maquiar.

(41) ?? Ela se tem de maquiar.

As intuições de um brasileiro a respeito dessas construções (que são mais correntes em Portugal) são algo incertas. Será necessário fazer um levantamento do uso dessas construções na língua escrita do Brasil, para obter um retrato fiel da situação.

8.1.5. Posposição do sujeito

A respeito da posposição do sujeito pouco resta a dizer, após as considerações feitas na seção 7.4. Sabemos que é um fenômeno bastante comum em português e que sofre restrições pouco compreendidas.

Em primeiro lugar, não se pode pospor o sujeito se houver objeto direto na oração:

(42) * Comeu uma pizza Sônia.

Comparar com

(43) Ontem chegou Sônia.

Em segundo lugar, a anteposição de certos termos parece favorecer a possibilidade de pospor o sujeito. Assim, não se diz

(44) * Está Sônia onde?

apesar de não haver problemas com

(45) Sônia está onde?

Mas, se antepusermos o elemento Q (*onde*), a posposição do sujeito se tornará possível:

(46) Onde está Sônia?

Analogamente, (43) ficaria bem pior se *hoje* não estivesse no início:

(47) ?? Chegou Sônia hoje.

O verbo da oração certamente exerce um efeito sobre a possibilidade de posposição do sujeito. Por exemplo, com *trabalhar* no lugar de *estar* (46) fica menos aceitável:

(48) ?? Onde trabalha Sônia?

Em certos casos, o verbo chega a causar franca inaceitabilidade; comparem-se

(43) Ontem chegou Sônia.

(49) * Ontem desmaiou Sônia.

Essas notas soltas ainda estão por ser integradas em uma análise coerente da posposição do sujeito. Por ora, teremos de contentar-nos com elas.

8.2. ORDEM DOS TERMOS NO SINTAGMA NOMINAL

Comparado com a oração, o sintagma nominal tem uma ordenação relativamente fixa, o que facilita sua descrição desse ponto de vista. A estrutura interna do SN já foi estudada em algum detalhe no capítulo 4; aqui vou apenas repetir a análise ali proposta.

O SN compreende nove posições fixas, estritamente ordenadas, a saber:

1ª: Determinante (Det);
2ª: Possessivo (Poss);
3ª: Reforço (Ref);
4ª: Quantificador (Qf);
5ª: Pré-núcleo externo (PNE);
6ª: Pré-núcleo interno (PNI);

7ª: Núcleo (NSN);

8ª: Modificador interno (ModI);

9ª: Modificador externo (ModE).

Além disso, há o numerador (Num), que pode colocar-se nos intervalos entre as cinco primeiras funções. Quaisquer aparentes mudanças de ordem (à parte as do numerador) são analisadas como mudança de função. Assim, em

(50) Um amigo meu

temos não um possessivo posposto, mas o item *meu* na função de modificador.

8.3. ESTRUTURAS ORACIONAIS DO PORTUGUÊS

8.3.1. Enumeração das estruturas

Uma das funções de uma gramática é enumerar, de alguma maneira, as estruturas — entre elas, as orações — possíveis na língua. Essa enumeração ainda não pode ser feita de maneira completa, pois depende de um levantamento muito completo da sintaxe da língua, levantamento de que não dispomos atualmente. Nesta seção, vou dar uma idéia de como se pode realizar essa tarefa, fornecendo ao mesmo tempo o que se pode dizer acerca das estruturas oracionais possíveis em português.

Comecemos considerando apenas três funções: sujeito, predicado e objeto direto. Vimos, na seção 8.1.1., que há uma regra que define as posições respectivas desses três termos:

RES-1: O ⟶ (Suj) + Pred + (OD)

Sabemos, além do mais, que as estruturas possíveis não são apenas resultado da aplicação dessa regra. Elas dependem igualmente de outras condições sintáticas: as relações de correspondência; as restrições de transitividade dos verbos; e outras restrições, como as que formam o sistema da concordância verbal estudado na seção 6.6.2.

Da interação de todos esses componentes, torna-se possível definir um número considerável de estruturas como bem formadas e outras tantas como mal formadas. Considerando apenas RES-1 e algumas correspondências totais, podemos prever muita coisa a respeito das orações possíveis em português.

Em primeiro lugar, as diferentes formas de aplicar RES-1 nos fornecem quatro alternativas, a saber:

(a) Suj + Pred + OD

(b) Suj + Pred

(c) Pred + OD

(d) Pred

Essas quatro estruturas são exemplificadas, respectivamente, pelas orações

(51) O gato mordeu Aninha.

(52) O gato fugiu.

(53) Mordi Aninha.

(54) Cheguei.

Levando em conta agora as relações de correspondência, podemos acrescentar as estruturas

(e) OD, + Suj + Pred (**Topicalização**)

(f) OD, + Pred (**Topicalização**)

(g) Pred + Suj (**Posposição de sujeito**)

(h) *ser* Suj *que* Pred + OD (**Clivagem**)

(i) *ser* OD *que* Suj + Pred (**Clivagem**)

(j) *quem* Pred + OD *ser* Suj (**Pseudo-clivagem**)

(l) *quem* Suj + Pred *ser* OD (**Pseudo-clivagem**)

A estrutura (e) é exemplificada por

(55) Aninha, o gato mordeu.

A estrutura (f) é geralmente de aceitabilidade algo baixa, devido provavelmente à ausência de sujeito. Isso deveria ser incluído na definição da topicalização:

(56) ?? Aninha, mordi.

As estruturas (g), (h), (i), (j) e (l) são exemplificadas respectivamente pelas orações

(57) Chegou um amigo meu.

(58) Foi o gato que mordeu Aninha.

(59) Foi Aninha que o gato mordeu.

(60) Quem mordeu Aninha foi o gato.

(61) Quem o gato mordeu foi Aninha.

Os mesmos mecanismos determinam uma série de estruturas que são mal formadas, justamente porque não há meio de gerá-las dentro do sistema. Por exemplo, não podemos ter * **Pred + Suj + OD**, porque essa ordenação não é autorizada por RES-1 e não há nenhuma relação de correspondência que permita colocar o predicado antes do sujeito. Essa previsão é confirmada:

(62) * Mordeu o gato Aninha.

Da mesma forma, podemos prever a inaceitabilidade de

* Pred + OD + Suj
* OD + Pred + Suj
* OD + Suj + Pred
* Suj + OD
* Suj
* OD
* *ser* Pred *que* Suj + OD

e várias outras. Essas seqüências ou não formam orações, ou são totalmente inaceitáveis. Alguns exemplos são:

(63) * O gato Aninha.

(64) * O gato.

(65) * Foi mordeu que o gato Aninha.

(64) é inaceitável enquanto oração, ou enquanto sujeito (de quê?). Naturalmente, a mesma seqüência

pode ser um sintagma nominal bem formado.

Como se vê, a partir de uma única regra, RES-1, mais vários outros dispositivos já estudados, podemos especificar uma boa fatia da sintaxe da oração portuguesa. Dizemos então que esses mecanismos **enumeram** o conjunto das estruturas oracionais possíveis na língua. Ampliando-se o sistema descritivo, com mais regras, regras mais abrangentes, novas relações de correspondência etc., deveremos chegar a uma enumeração completa das estruturas sintáticas do português.

8.3.2. A sintaxe dentro da descrição geral

Nunca é demais relembrar que esse mecanismo, mesmo quando completo, não será suficiente para definir todas as possibilidades de expressão. O mecanismo enumera apenas as estruturas **sintaticamente** bem formadas — mas as exigências da língua vão bem além da boa formação sintática. Uma estrutura só é aceitável quando é bem formada sob todos os pontos de vista: fonológico, morfológico, sintático, semântico e pragmático. Vou dar abaixo um exemplo de violação de cada um desses componentes. Em primeiro lugar, suponhamos que alguém diga

(66) O gato mordeu Aninha.

pronunciando todos os **oo** bem abertos; algo como

(66') * ó gató mórdeu Aninha.

A frase será rejeitada porque está pronunciada incorretamente, ou seja, porque viola as regras do componente fonológico. O fato de ela ser sintaticamente bem formada não a salvará: uma expressão só é plenamente aceitável se está de acordo com as exigências de todos os componentes da gramática.

Agora consideremos

(67) * O gato mordou Aninha.

Aqui o problema é morfológico: o verbo *morder* foi conjugado incorretamente, como se fosse da primeira conjugação. Como resultado, a oração é inaceitável.

A oração pode ser sintaticamente malformada, como no exemplo já visto

(62) * Mordeu o gato Aninha.

A pronúncia pode estar correta e a morfologia perfeita; como há violação sintática, a frase é rejeitada.

Podemos ter também violação de caráter semântico; por exemplo,

(68) * O gato transcorreu tranqüilamente.

O problema não está na fonologia, nem na sintaxe, nem na morfologia, mas na semântica: a idéia de "transcorrer" só se aplica a períodos de tempo (*a semana transcor-*

reu tranqüilamente), ou a eventos que se desenrolam no tempo (*a reunião transcorreu tranqüilamente*); dizer que *o gato transcorreu* pode estar sintaticamente correto, mas não faz sentido. Por isso (68) é também inaceitável.

Finalmente, pode haver problemas de ordem pragmática, como quando alguém está sozinho no elevador com um amigo e diz

(69) Vocês todas estão irritadas comigo?

Do ponto de vista estritamente gramatical não há nada de errado com (69). Mas podemos dizer que, naquele contexto, a frase é inaceitável. O falante usou uma estrutura correta na hora errada, por assim dizer.

8.3.3. Ampliando a análise

Até agora só levamos em conta três termos da oração: o sujeito, o predicado e o objeto direto. Naturalmente, será necessário ampliar a análise para dar conta do posicionamento dos demais termos. Isso está por ser feito em detalhe; mas já é possível dar uma idéia de como será a imagem final da oração, no que diz respeito à ordem de seus termos.

Para efeitos da descrição, vamos adotar como básica a seguinte estrutura para a oração em português:

Suj + Pred + OD

Essa estrutura é gerada pela regra que chamamos RES-1. Sabemos que ela pode ser modificada em várias estruturas correspondentes (seção 8.3.1.). E veremos agora que pode ser acrescida de diversos termos; as posições desses termos serão definidas em função da ordenação básica fornecida pela estrutura acima.

Em primeiro lugar, vejamos o complemento do predicado. Esse termo ocorre no **mesmo** lugar do objeto direto, o que significa que os dois se excluem mutuamente. E, com efeito, nunca se encontra objeto direto e complemento do predicado na mesma oração. Vamos exprimir isso modificando a formulação da RES-1 da seguinte maneira:

RES-1: O ⟶ (Suj) + Pred + (OD/CP)

(onde a barra inclinada indica as alternativas disponíveis para preencher o terceiro termo: escolhe-se um dos dois, mas nunca ambos).

Desse modo, fica incluído o CP dentro da estrutura da sentença em português.

Agora tomemos os seguintes termos: adjunto oracional (AO), atributo (Atr), adjunto adverbial (AA) e predicativo (Pv). Esses elementos se caracterizam por ocorre-

rem nos intervalos entre as três posições da estrutura básica. Para facilitar a referência, vamos numerar esses intervalos, definindo quatro posições, da seguinte forma:

1 Suj **2** Pred **3** OD/CP **4**

Os termos em exame se colocam em uma ou mais das posições acima marcadas pelos números 1 a 4, a saber:

AO: ocorre nas posições 1, 2, 3, 4;
Atr: ocorre nas posições 1, 2, 3, 4;
AA: ocorre nas posições 3, 4;
Pv: ocorre nas posições 3, 4.

Os exemplos são fáceis de encontrar; vou dar apenas um de cada função. Com AO, temos:

(70) a. Infelizmente, o gato mordeu Aninha.
b. O gato, infelizmente, mordeu Aninha.
c. O gato mordeu, infelizmente, Aninha.
d. O gato mordeu Aninha, infelizmente.

Com Atr:

(71) a. Todo dia o gato morde Aninha.
b. O gato todo dia morde Aninha.
c. O gato morde todo dia Aninha.

Com AA:

(72) a. * Totalmente, Aninha decepcionou Cláudio.
b. * Aninha, totalmente, decepcionou Cláudio.
c. Aninha decepcionou totalmente Cláudio.
d. Aninha decepcionou Cláudio totalmente.

E com Pv:

(73) a. * Um amigo, Aninha considera o gato.
b. * Aninha um amigo considera o gato.
c. Aninha considera um amigo o gato.
d. Aninha considera o gato um amigo.

Resta, das funções de nível oracional, o adjunto circunstancial (AC). Este parece ocorrer nas posições 1, 3 e 4 — mas os dados não são muito consistentes, o que possivelmente decorre de o AC representar mais de uma função. Até que isso seja devidamente investigado, é melhor adiar a questão da ordem do AC dentro da oração.

O que se viu acima deve servir para ilustrar como se pode pesquisar a estrutura interna da oração em português, no que concerne à ordem dos termos. Nem é preciso dizer que os resultados expostos devem ser tomados como provisórios.

TERCEIRA PARTE: SEMÂNTICA

9 Semântica: preliminares

9.1. SEMÂNTICA E PRAGMÁTICA

O primeiro problema a enfrentar ao se encetar o estudo do significado das formas lingüísticas é o de distinguir qual a porção do significado que é oriunda da interpretação das estruturas e dos itens léxicos e qual a porção que provém do conhecimento que o falante tem de fatores extralingüísticos, tais como: a situação de comunicação; as relações existentes entre os interlocutores; o conhecimento prévio que eles possam ter do assunto, e assim por diante.

Para tomar um exemplo quase clássico, vejamos o caso das perguntas. Pode-se fazer uma pergunta utilizando uma oração interrogativa, como

(1) Que horas são?

Mas o mesmo pedido de informação poderia ser formulado sem utilizar uma interrogativa. Se estou em uma festa, viro-me para um amigo e digo

(2) Eu não sei que horas são.

ele provavelmente interpretará essa oração declarativa como um pedido de informação equivalente à interrogativa (1).

Outra maneira de perguntar as horas seria

(3) Você sabe que horas são?

Essa é uma maneira muito comum de se perguntar as horas. Mas é interessante observar que (3), estritamente falando, não tem a estrutura de um pedido de informação sobre as horas; antes, é um pedido de informação sobre se o interlocutor **sabe** quantas horas são. A resposta poderia ser um simples "sim", e em algumas situações essa resposta seria adequada.

A conclusão é bem clara: nossa compreensão dos enunciados não é função exclusiva de um processamento das estruturas lingüísticas contidas neles. É também parcialmente função de nossa percepção da situação em que nos encontramos, com quem nos estamos comunicando, aquilo que sabemos e aquilo que acreditamos que nosso interlocutor

também saiba. É como se o processo de interpretação compreendesse duas etapas: primeiro, temos um sistema que extrai do enunciado aquilo que é possível depreender a partir somente da estrutura formal (morfossintática). O resultado é uma representação semântica, que poderíamos chamar o **significado literal** do enunciado. Em um segundo momento, essa representação semântica se associa a uma série de outros fatores, ligados ao contexto da comunicação e ao conhecimento prévio existente (ou pressuposto como tal) na memória do falante e do ouvinte. Resulta daí algo que, por falta de termo melhor, chamarei o **significado final**.

Voltando ao exemplo (3), a representação semântica corresponde ao significado literal "pergunta que espera resposta sim ou não": o falante deseja saber se o ouvinte sabe as horas. Se a pergunta for compreendida literalmente, uma resposta adequada será "sim".

Mas nem sempre os enunciados são compreendidos literalmente, porque há fatores contextuais (extralingüísticos) que também são levados em conta. Ao ouvir (3), o interlocutor pode raciocinar do seguinte modo:

(a) ele está sem relógio;

(b) já está ficando tarde, e provavelmente ele tem hora para sair desta festa;

(c) a informação de que "eu sei que horas são", em si, não seria de nenhuma utilidade para ele, no atual contexto;

(d) logo, ele não está querendo saber se eu sei as horas; deve querer saber que horas são.

Após esse raciocínio (implícito, evidentemente), a resposta mais provável será algo como

(4) São quase quatro horas.

Mais um exemplo: em uma frase como

(5) Roberto procurou o médico porque disseram que ele estava abatido.

a palavra *ele* pode, em princípio, referir-se tanto a Roberto quanto ao médico. No entanto, normalmente a entendemos como referindo-se a Roberto, porque em geral o que acontece é o abatido ir ao médico, e não o paciente procurar um médico abatido. Note-se como a referência de *ele* passa a ser o médico se mudamos a segunda oração:

(6) Roberto procurou o médico porque disseram que ele era o melhor da cidade.

Nos casos de (5) e (6), a estrutura é ambígua e admite o relacionamento de *ele* com *Roberto* ou então com *o médico.* Mas nosso conhecimento extralingüístico (de coisas como: pessoas abatidas procuram o médico; ao procurar um médico, dá-se preferência àqueles que têm boa reputação profissional) desfaz a ambigüidade, de modo que, em situações normais, as frases (5) e (6) não são sentidas como ambíguas.

Há casos em que a ambigüidade não existe, porque a referência do pronome é determinada sem margem de dúvida. Nesses casos, a representação semântica (significado literal) não pode ser alterada por fatores de natureza pragmática. Um exemplo é a referência do pronome *o* na frase

(7) Roberto mandou que Caio o penteasse.

Só é possível entender que *o* se refere a Roberto, nunca a Caio. Mas se no lugar de *o* colocarmos *se*, o pronome terá de ser entendido obrigatoriamente como referindo-se a Caio, e não a Roberto:

(8) Roberto mandou que Caio se penteasse.

Como se vê, em muitos casos é como se o significado literal "viesse à tona", sem interferência de fatores pragmáticos; em outros, o significado literal não é mais que um estágio intermediário na computação do significado final.

O primeiro componente da depreensão do significado (o que produz o significado literal) denomina-se **semântica**; o segundo, que computa o significado final com base no significado literal, mais fatores extralingüísticos, chama-se **pragmática**. Neste trabalho, estou pressupondo que a semântica e a pragmática podem ser descritas como componentes separados (ao lado da sintaxe, da morfologia, da fonologia etc.). No presente capítulo, trataremos apenas da semântica, o que significa que evitaremos questões relativas ao efeito do contexto na computação do significado. Já o estudo da pragmática está fora do escopo desta *Gramática*.

Devo alertar o leitor que a hipótese da separação da semântica e da pragmática em componentes distintos não é universalmente aceita. Na verdade, trata-se de um dos grandes pontos de discussão da lingüística atual, e muitos negam a possibilidade de separar os dois tipos de fatores. Aqui, como minha tarefa é a descrição do português, sou forçado a tomar partido e optei pela alternativa que me parece mais plausível: uma decisão ditada mais pela necessidade prática do que por uma convicção profunda.

Adotaremos, portanto, a teoria segundo a qual as formas lingüísticas (orações, por exemplo) são interpretadas por um componente semântico. Esse componente é composto, como os outros, por regras (chamadas **regras de interpretação semântica**, ou simplesmente **regras de interpretação**), mais outros mecanismos, tais como filtros e restrições várias. O resultado da aplicação das regras de interpretação às formas lingüísticas é uma representação semântica, que corresponde ao significado literal. Como vimos, a representação semântica poderá ser ainda alterada por razões de pragmática, antes que se obtenha o significado final.

A partir do capítulo 10, apresentarei alguns exemplos do funcionamento desse sistema. Os exemplos não chegam, nem de longe, a cobrir a estrutura semântica da língua (que, aliás, é pouquíssimo conhecida em seus traços gerais). Aqui, como em tantos outros pontos desta *Gramática*, tenho que limitar-me a apontar direções de pesquisa.

9.2. REGRAS, TRAÇOS E REPRESENTAÇÕES SEMÂNTICAS

A descrição da semântica de uma língua apresenta dois aspectos principais: a semântica dos itens léxicos e a semântica das formas gramaticais. A primeira se ocupa do significado individual dos itens léxicos, isto é, da descrição dos traços semânticos que caracterizam o significado de cada um deles. Esse estudo não se coloca propriamente na gramática, ligando-se antes ao estudo do léxico (lexicologia); por isso, aqui não entraremos em pormenores no que diz respeito às matrizes de traços semânticos que integram os itens léxicos. Por outro lado, como se verá, estas tampouco podem ser ignoradas, porque constituem uma contribuição fundamental para a interpretação das estruturas; teremos, pois, de dar-lhes alguma atenção, ainda que informal e sumária.

A semântica das formas gramaticais descreve a contribuição da estrutura morfossintática à interpretação semântica. Esse pode ser considerado um estudo gramatical; leva em consideração questões tais como:

— Como se interpreta o modificador de um NSN?

— Que significado tem o sujeito de uma oração?

— Como se identifica o agente de uma ação em uma oração sem sujeito?

e assim por diante.

A semântica das formas gramaticais é descrita através de regras semânticas, que se aplicam a estruturas morfossintáticas e lhes atribuem uma interpretação; essas regras serão estudadas e exemplificadas nos capítulos 10 e 11. Começarei a exposição dando uma rápida idéia de como se organiza a matriz semântica dos itens léxicos, para logo passar às regras que se aplicam às estruturas.

Os itens léxicos que compõem as sentenças da língua têm um significado próprio: *camelo* não significa a mesma coisa que *cachimbo*. Esse significado é codificado no item através de **traços semânticos**, dos quais evidentemente existe uma variedade imensa. Por exemplo, tanto *camelo* quanto *cachimbo* são identificados como objetos concretos (que podem ser pesados, tocados etc.); mas só *camelo* é identificado como um ser animado: isso quer dizer que ambas as palavras têm o

traço [+concreto], mas só *camelo* tem o traço [+animado].

Além desses traços, que caracterizam o significado próprio de cada item, há outros que expressam condições de compatibilidade de cada item com os demais em uma mesma sentença. Assim, o verbo *quebrar*, como veremos, exige a presença de um paciente (elemento que sofre a ação) na representação semântica de qualquer frase em que figure. Não se trata de uma exigência sintática, porque esse paciente pode ser representado na sentença como objeto direto ou como sujeito (ver exemplificação na seção 10.2.). E existem também traços que estipulam relações de correferência entre diversos elementos da representação semântica; por exemplo, o verbo *mandar* exige que a pessoa que recebe a ordem (a **meta**) seja a mesma que pratica a ação expressa pela oração subordinada (o **agente**).

Uma oração (ou qualquer forma lingüística) é composta de itens léxicos relacionados através de funções sintáticas (ou morfológicas, no caso de palavras). E cada item léxico encerra, como vimos, uma matriz de traços semânticos. Mas o significado da oração não é a soma dos traços semânticos dos diferentes itens que a compõem. Se fosse assim, as frases abaixo seriam sinônimas, já que se compõem dos mesmos itens léxicos:

(9) Patrícia insultou Bernardo.

(10) Bernardo insultou Patrícia.

Acontece que as funções sintáticas também contribuem para o significado das sentenças. A diferença entre (9) e (10) é que as funções sintáticas de *Patrícia* e *Bernardo* estão trocadas; o resultado é uma nítida diferença de significado. Por isso, necessitamos de regras de interpretação para atribuir traços de significado aos constituintes das sentenças de acordo com a função de cada um (não é só isso o que as regras de interpretação fazem; mas vamos limitar-nos à interpretação das funções, para simplificar esta exposição inicial).

Tomando, como exemplo, a frase

(9) Patrícia insultou Bernardo.

podemos resumir o processo de interpretação assim:

(a) os traços semânticos dos itens léxicos nos fornecem informações tais como: "*Patrícia* é o nome de um ser humano, feminino"; "*Bernardo* é o nome de um ser humano, masculino"; "*insultar* é uma ação" etc.;

(b) as regras semânticas especificam o sujeito *Patrícia* como sendo o agente da ação, e o objeto direto *Bernardo* como sendo o paciente.

O resultado da aplicação das regras semânticas à estrutura da sentença (9), mais os traços semânticos associados a cada um dos itens, é a **representação semântica** dessa sentença. A representação semântica da

frase (9) explicita, entre outras informações, que Patrícia praticou uma ação, que essa ação foi sofrida por Bernardo, que a ação foi a de insultar, que foi praticada no passado etc.

Acabo de dar uma noção muito sumária do que vêm a ser regras de interpretação, traços semânticos e representações semânticas. A partir do capítulo 10, veremos alguns exemplos da aplicação das regras na derivação de representações semânticas; esses exemplos devem servir para deixar bem claro como funciona o componente semântico da língua.

9.3. SUMÁRIO: FUNCIONAMENTO DO COMPONENTE SEMÂNTICO

Entende-se, portanto, o componente semântico da gramática como um conjunto de regras que atribuem às estruturas certas representações semânticas. Nos capítulos seguintes, veremos diversos exemplos de regras semânticas, que deverão deixar mais clara essa noção.

Além disso, cada item léxico compreende, ao lado de suas matrizes de traços fonológicos, sintáticos e morfológicos, também traços semânticos. Por exemplo, seja o item léxico *beber*[1]. O item completo inclui uma matriz fonológica, que contém as instruções para a pronúncia do item; depois, uma matriz que estabelece suas propriedades morfológicas (por exemplo, trata-se de uma forma que admite variação de pessoa e número); depois, uma matriz sintática, que estipula que *beber* pode ser núcleo de um sintagma nominal, não pode ser modificador etc. Finalmente, temos a matriz semântica, que singulariza o item no que diz respeito a seu significado.

A matriz semântica compreende alguns traços bastante gerais, isto é, que também se aplicam a muitos outros itens, como "ação", "evento"; outros traços que são próprios desse item ou de alguns poucos itens, como o que informa que se trata de uma ação relativa à ingestão de algum líquido; e ainda traços de compatibilidade, que de algum modo descrevem o tipo de agente e paciente que *beber* admite (por exemplo, um gato bebe, mas um muro não; e pode-se beber água, ou mesmo tinta, mas não farinha ou pastéis).

Essa descrição da matriz semântica dos itens léxicos é bastante vaga, mas terá que servir para nossos objetivos. Não seria possível entrar em grandes detalhes a respeito da estrutura das matrizes semânticas sem nos desviarmos muito do nosso assunto; o leitor encontrará uma descrição um pouco mais detalhada no capítulo 14.

---

[1] O léxico não é composto apenas de palavras; certamente, inclui ainda uma lista de morfemas presos (que só aparecem como partes de palavras), como *-s* "plural", ou o sufixo *-mente,* que se encontra em *rapidamente* etc. Estou tomando exemplos de palavras para facilitar a exposição (ver no capítulo 14 uma descrição da estrutura do léxico).

Assim compreendido, o componente semântico tem a função de derivar representações semânticas a partir dos itens léxicos e das funções sintáticas. No exemplo

(9) Patrícia insultou Bernardo.

os itens léxicos fornecem parte da representação semântica (os significados individuais de *Patrícia, Bernardo, insult-,-ou*). O restante é derivado das funções sintáticas (*Patrícia* é o sujeito, e portanto, neste caso, o agente etc.).

Essa representação, em certos casos, pode ser mal formada — isto é, pode "não fazer sentido". Por exemplo, a frase

(11) * Meu vizinho bebeu todo o papel.

é sintática e morfologicamente bem formada, mas alguma coisa está errada com ela. É que a representação semântica que se deriva de (11) desobedece a uma das condições do significado de *beber*, a que estabelece que o paciente da ação de beber deve ser um líquido (e a matriz semântica do item *papel* o descreve como sólido). Conseqüentemente, a frase é marcada como mal formada.

Conforme se vê, as regras semânticas, como as demais, podem funcionar como uma espécie de filtro, bloqueando frases que apresentam alguma má formação. Uma frase plenamente aceitável é aquela que passou com sucesso os filtros em todos os níveis.

É interessante observar que frases semanticamente malformadas podem ser utilizadas em contextos particulares: na chamada linguagem figurada, em poemas, em fábulas etc. Nesses contextos, as violações semânticas são utilizadas funcionalmente, para causar certos efeitos estéticos, ou para ajudar a caracterizar um ambiente fora do normal etc. Note-se que violações sintáticas ou morfológicas quase nunca são utilizadas dessa maneira.

9.4. ALGUMAS NOÇÕES BÁSICAS

Nesta seção, examinaremos algumas noções indispensáveis ao estudo da semântica. Nem a lista das noções tratadas nem seu tratamento têm a pretensão de ser completos; o leitor interessado na semântica geral ou portuguesa deverá necessariamente fazer um estudo cuidadoso de obras de introdução à semântica. Como primeiro passo, sugiro a leitura de Ilari & Geraldi, 1985.

9.4.1. Sinonímia, antonímia

Chamam-se **sinônimos**

> palavras de sentido igual ou aproximado. [Cegalla, 1987, p. 261]

Essa noção se estende a frases e sintagmas, que também podem ter sen-

tido semelhante e, portanto, podem igualmente ser sinônimos.

Mas que quer dizer, exatamente, "sentido igual ou aproximado"? Para tentar esclarecer a questão, vejamos primeiro exemplos de frases. As duas frases abaixo podem ser consideradas sinônimas:

(12) a. As formigas comeram o bolo.

b. O bolo, as formigas comeram.

As duas, na verdade, não são exatamente equivalentes de todos os pontos de vista. Assim, em (12b), o falante parece estar falando principalmente do bolo, ou seja, tomando o bolo como **tópico**, ou assunto principal; em (12a), essa seleção do bolo como tópico não existe.

Mas (12a) e (12b) são equivalentes de um ponto de vista importante: ambas têm o mesmo valor de verdade, em qualquer circunstância imaginável. Ou seja, não é possível imaginar uma situação em que (12a) seja verdadeira e (12b) falsa, ou vice-versa. Para usar um termo emprestado da lógica, diremos que as duas frases **implicam-se mutuamente**. O que devemos perguntar agora é: será que a implicação mútua é suficiente para definir a sinonímia?

Pode-se aplicar a mesma noção a palavras isoladas; assim, podemos dizer que tudo aquilo que é um "costume" é também um "hábito", e vice-versa; essas palavras também se implicam mutuamente (creio que aqui estamos estendendo um pouco a noção lógica de implicação mútua; mas a noção assim estendida nos servirá no contexto desta discussão).

A definição de "sinônimos" como "formas diferentes, com significado tal que se implicam mutuamente", entretanto, esbarra com certas dificuldades. Vejamos o seguinte par de frases:

(13) a. Pedro é o marido de Laura.

b. Laura é a esposa de Pedro.

Acontece que essas frases se implicam mutuamente — não há possibilidade de Pedro ser o marido de Laura sem que Laura seja a esposa de Pedro. No entanto, a maioria das pessoas negará que as frases sejam sinônimas.

Ou então tomemos o exemplo das palavras *costume* e *hábito*: vimos que seus significados se implicam mutuamente. No entanto, a língua impõe certas limitações ao emprego dessas palavras, e essas limitações não são idênticas, de tal forma que nem sempre se pode substituir *hábito* por *costume* ou vice-versa. Por exemplo, podemos falar dos *usos e costumes da nossa tribo*, mas não dos * *usos e hábitos da nossa tribo*. Ou então, podemos dizer que *fumar no elevador é um mau hábito*, ao passo que ? *fumar no elevador é um mau costume* é pelo menos um tanto estranho.

Em outras palavras, a noção (lógica) de implicação mútua não pode ser utilizada (pelo menos sozinha) para definir a noção léxica de

"sinonímia". A implicação mútua pode resultar de fatores não-lingüísticos (como nosso conhecimento de que todo marido tem uma esposa); e duas palavras que se implicam mutuamente podem, ainda assim, ocorrer em situações diferentes, devido a imposições de uso que não têm nada a ver com o significado.

Na verdade, ainda não se conseguiu uma definição precisa do que vem a ser "sinonímia"; isso é geralmente reconhecido pelos gramáticos e lexicógrafos, que sempre afirmaram que sinônimos perfeitos são uma raridade. A noção de sinonímia permanece intuitiva e bastante nebulosa.

Dito isso, é preciso reconhecer que os dicionários de sinônimos são muito úteis. Sua utilidade provém do fato de que nos fornecem grupos de palavras que, embora não sejam sinônimos perfeitos, são de significado muito próximo. São, na verdade, antes de tudo dicionários de idéias afins (ou melhor, de palavras que veiculam idéias afins); como disse, com toda a razão, um dos nossos lexicógrafos,

> [...] a um dicionário como este mais propriamente se deveria chamar ANALÓGICO. [Fernandes, 1944, p. 7]

A noção de **antonímia** é ainda menos nítida do que a de sinonímia. Aqui é impossível encontrar mesmo a coerência aproximada que se percebe no caso da sinonímia — e, correspondentemente, as listas de antônimos são bem menos úteis do que as de sinônimos, por serem menos confiáveis. Vejamos alguns poucos exemplos; encontrei os seguintes pares de antônimos em gramáticas tradicionais:

normal — anormal
velho — novo
bom — mau
emigrante — imigrante
pré-nupcial — pós-nupcial.

Não há nenhuma relação semântica (e/ou lógica) que se aplique a todos esses casos.

No caso de *normal — anormal,* temos duas qualidades que se excluem mutuamente, sem possibilidade de meio-termo: alguma coisa é normal ou então é anormal, e a separação é nítida.

Já em *velho — novo* temos uma relação diferente: um objeto (ou pessoa) não pode ser *velho* e *novo* ao mesmo tempo; mas a separação não é nítida, porque há uma infinidade de graus intermediários. Há simplesmente coisas e pessoas que não são claramente novas nem velhas.

O caso de *bom — mau* é semelhante ao de *velho — novo,* mas com a complicação de que a diferenciação entre as duas qualidades depende de uma avaliação subjetiva, que pode variar de pessoa para pessoa. Assim, podemos dizer com segurança que um homem de 88 anos é velho; mas classificar um prato de frango com quiabo como "bom" ou

"mau" depende inteiramente de quem faz o julgamento.

A dupla *emigrante* — *imigrante* nos leva para um campo semântico totalmente diferente. Aqui se trata, em última análise, da direção de um movimento. Além disso, todo emigrante é também imigrante: quem sai do Japão para vir viver no Brasil é emigrante para os japoneses, mas imigrante para os brasileiros. Não vejo nenhuma proximidade entre essa oposição e as oposições examinadas acima.

Finalmente, a diferença entre *pré-nupcial* e *pós-nupcial* é uma diferença de momento de realização de um ato (uma cerimônia, um exame médico etc.). Novamente, temos aqui uma oposição que não se identifica, nem sequer se aproxima das demais.

Apesar dessas diferenças gritantes, os cinco pares de palavras examinados são considerados pares de antônimos. Aqui, a situação é de tal heterogeneidade que me parece conveniente abandonar, pelo menos por ora, a noção de "antonímia". Enquanto não se obtiver uma conceituação muito melhor do que a atual, essa será uma noção sem utilidade nenhuma.

9.4.2. Homonímia e polissemia

Existem algumas noções básicas que, não obstante sua importância, nunca foram convenientemente definidas. Um dos exemplos mais gritantes é o da diferença entre **homonímia** e **polissemia**. Como veremos, distinguir claramente entre casos de homonímia e casos de polissemia é fundamental para se conferir rigor à descrição gramatical; no entanto, e apesar de muitos esforços nesse sentido, até hoje não se conhece uma maneira de distinguir claramente esses dois fenômenos.

Vejamos primeiro do que se trata. Sabemos que uma palavra pode ter vários significados. Assim, *verde* pode ser o nome de uma cor, ou então um estágio na maturação de uma fruta; tanto é assim que se pode dizer de uma fruta amarela que ainda está verde.

O problema que se coloca é o seguinte: devemos considerar *verde* uma única palavra ou duas? Se a considerarmos uma única palavra, teremos de dizer que tem mais de um significado, ou seja, que é **polissêmica**. Já se distinguirmos duas palavras *verde*, diremos das duas que têm a mesma pronúncia e grafia, e que são **homônimas**. Esse é um dilema que até hoje não recebeu solução satisfatória.

A saída tradicional se baseia em um conjunto heterogêneo de critérios. Segundo essa proposta (exemplificada em Dubois & Dubois, 1971), distinguem-se duas palavras (dois itens léxicos) quando há uma diferença de classe gramatical, ou então quando há uma diferença semântica grande e nítida.

Isso permite uma decisão bem fácil em certos casos; por exemplo, *canto* pode ser uma forma do verbo *cantar*, como em

(14) Eu só canto samba.

ou um substantivo, como em

(15) Interpretamos um canto folclórico.

Aqui, a diferença de classe gramatical nos levará a distinguir duas palavras; trata-se, pois, de um caso de homonímia.

Já o critério da diferença semântica é muito mais problemático. Assim, no caso de *manga* "fruta" e *manga* "parte da vestimenta", provavelmente a distância entre as duas acepções é suficiente para que se distingam duas palavras. Mas que fazer nos muitíssimos casos intermediários? Como distinguir com segurança uma diferença semântica "grande" de uma "pequena"? Para se ter uma idéia do problema, comparem-se os seguintes casos:

(16) a. *fio* (de linha) — *fio* (elétrico)

b. *pintar* (um quadro) — *pintar* (uma parede)

c. *choque* (elétrico) — *choque* (susto)

d. *roda* (de carroça) — *roda* (de amigos)

e. *batida* (trombada) — *batida* (bebida)

f. *papel* (para escrever) — *papel* (de um ator teatral)

g. *mosca* (inseto) — *mosca* (centro do alvo)

h. *pena* (de ave) — *pena* (castigo).

Em todos esses casos, a diferença de significado é evidente. No exemplo (a), a diferença pode ser considerada "pequena" e, no exemplo (h), é claramente "grande". Mas onde está o limite entre as diferenças "grandes" e as "pequenas"? Se não tivermos um critério facilmente aplicável para estabelecer esse limite, a proposta tradicional para distinguir homonímia de polissemia não terá grande utilidade.

O problema vem de que a polissemia não é, como se poderia pensar, um defeito das línguas (que insistiriam em dar o mesmo nome a coisas diferentes). A polissemia é uma propriedade fundamental das línguas humanas, que sem ela não poderiam funcionar eficientemente. Seria impraticável dar um nome separado a cada "coisa", incluindo aquelas que nunca vimos. Ao nos depararmos com um objeto nunca visto antes — digamos, um novo modelo de bicicleta — ficaríamos sem recursos para denominá-lo. Mas não é assim que a linguagem e a mente trabalham. Ao encontrar um objeto novo, tentamos imediatamente "reconhecê-lo", encaixando-o em alguma categoria já existente na memória (e na língua). Ao vermos um animal desconhecido, em geral tendemos a chamá-lo pelo nome de um animal já conhecido; assim, chamamos *formiga* aos representantes de milhares de espécies diferentes de

insetos; e assim uma criança diz *"cocó"* ao ver pela primeira vez uma avestruz. A polissemia confere às línguas humanas a flexibilidade de que elas precisam para exprimirem todos os inumeráveis aspectos da realidade.

Conseqüentemente, a maioria das palavras são polissêmicas em algum grau. Palavras não-polissêmicas são raras e freqüentemente são criações artificiais, como os termos técnicos das ciências: *fonema, hidrogênio, pâncreas* etc. Nestes casos, a polissemia é realmente um inconveniente; mas o discurso científico, em sua procura de univocidade semântica, difere enormemente da fala normal das pessoas. Nesta, a polissemia é indispensável.

Como se vê, o problema é não apenas de difícil solução, mas também bastante grave. Como elaborar uma descrição rigorosa da estrutura de uma língua se não temos um critério confiável que nos diga exatamente onde termina um item léxico e onde começa outro? Confrontados com os exemplos abaixo:

(17) a. Achei um rato nessa gaveta. [= "encontrei"]

b. Achei que a festa já tinha acabado. [= "pensei"]

c. Acho Marta uma boa aluna. [= "considero"]

como poderemos dizer se temos aí um verbo (com três significados e três tipos de transitividade distintos), ou três (cada um com seu significado e sua transitividade própria), ou ainda dois [digamos, agrupando o de (b) com o de (c)]?

A dificuldade com o critério semântico provém de que ele depende de se delimitar os conceitos veiculados pelos itens léxicos para daí delimitar os próprios itens. E a delimitação dos conceitos é um problema ainda sem solução conveniente. De qualquer modo, muito trabalho se está realizando nesta área: alguns procuram uma maneira puramente formal (não-semântica) de delimitar os itens léxicos, enquanto outros tentam aperfeiçoar nossos recursos de delimitação de conceitos. Como se trata de questão em aberto, não temos alternativa senão procurar fazer o melhor uso da definição disponível, evitando, sempre que possível, as armadilhas piores. Não é um processo rigorosamente científico, mas é o que todos os lingüistas adotam, por falta de coisa melhor.

9.4.3. Tempo, aspecto, modo

Vamos agora examinar três categorias semânticas que são expressas pelo verbo: o **tempo**, o **aspecto** e o **modo**.

Antes de começar, será necessário alertar o leitor para o fato de que a nomenclatura tradicional é aqui especialmente enganadora. Fala-se, tradicionalmente, de "tempos" e "modos" verbais, e os nomes de

certos tempos verbais incluem termos como "perfeito", "imperfeito", que são em geral usados para designar aspectos. Além disso, a nomenclatura tradicional não distingue o lado formal do lado semântico do fenômeno; e, na área de tempos, aspectos e modos, a discrepância entre forma e significado é muito grande.

Exemplos dessa discrepância são: (a) o tempo chamado "presente do indicativo" nem sempre se refere a um fato que se está dando no momento presente; (b) nenhuma categoria semântica conhecida é expressa com exclusividade pelas formas do "modo subjuntivo"; (c) o aspecto imperfectivo pode ser veiculado por muitas formas que não são chamadas "imperfeito".

Para facilitar a leitura, optei por manter as designações tradicionais para as **formas** verbais. Assim, a forma *faço* continuará a chamar-se "presente do indicativo"; mas isso não quer dizer que *faço* exprima sempre um fato presente (podemos dizer, por exemplo, *amanhã eu faço isso para você*). Estaremos, assim, distinguindo o **tempo verbal**, categoria morfológica (portanto formal), do **tempo semântico**, categoria de significado. Em *amanhã eu faço isso para você*, a forma *faço* exemplifica o tempo verbal chamado "presente (do indicativo)", mas veicula o tempo semântico "futuro" — isto é, exprime um fato a se realizar em algum momento por vir. A forma *faço* pode exprimir também eventos que não são futuros, como em *eu faço tapeçarias para vender*. Em resumo, a vinculação entre tempo verbal (morfológico) e tempo semântico é complexa; e é uma pena que tenhamos de referir-nos a ambos com a mesma palavra. Espero que a qualificação "verbal" e "semântico" evite confusões.

9.4.3.1. Tempo

Sabemos que as frases portuguesas se entendem em geral como localizadas em determinado momento no tempo. Assim, faz diferença dizer

(18) Manuel está picando a couve.

ou

(19) Manuel picou a couve.

Essas duas frases diferem semanticamente de várias maneiras; uma dessas é que com (18) entendemos que a ação de "picar couve", praticada por Manuel, está se desenvolvendo no momento em que se fala, e com (19) entendemos que a ação se deu em algum momento anterior. Dizemos então que há uma diferença de **tempo** entre (18) e (19).

Antes de prosseguir, é necessário fazer umas observações. Vimos acima que a noção semântica de "tempo" não se identifica, nem sequer se relaciona de maneira simples, com a categoria morfológica "tempo (verbal)". Vejamos agora mais uns exemplos.

Tome-se o presente do indicativo nas frases

(20) Manuel pica a couve neste momento.

(21) A água ferve a 100 graus.

(22) Em 1822, o Brasil se torna politicamente independente.

(23) Amanhã bem cedo eu termino este serviço.

Todas as quatro frases têm verbo no presente do indicativo; no entanto, a referência temporal (o tempo semântico) é diferente em cada uma. Em (20) — aliás, uma frase um tanto marginal para o falante brasileiro — a ação de "picar" se entende como realmente contemporânea ao ato de fala. Em (21), o fato mencionado é independente de tempo: trata-se de uma propriedade da água, válida em qualquer momento. Em (22) (um exemplo de "presente histórico"), o fato se deu no passado. Em (23), ainda vai se realizar.

O presente do indicativo é a forma verbal mais versátil do ponto de vista da referência temporal; com outros tempos, as possibilidades são muito mais restritas. O futuro do indicativo, por exemplo, só pode ser usado para expressar fatos ainda à ocorrer. O pretérito perfeito nunca pode exprimir fatos presentes; mas pode exprimir fatos passados ou futuros — neste último caso, exprime um fato passado por referência a outro fato, estando ambos no futuro:

(24) Amanhã quando você chegar eu já limpei tudo.

Tanto *chegar* quanto *limpei* exprimem fatos a se realizarem no futuro.

A segunda observação a ser feita é que o tempo semântico não se confunde com o tempo cronológico que é registrado nos relógios e nos calendários. Em uma narração colocada no passado, por exemplo, uma frase com tempo semântico "presente" está, em relação ao leitor, no passado, como em

(25) Manuel disse: "Estou picando a couve".

Naturalmente, *estou picando* é presente semântico em relação ao momento da fala de Manuel, mas é passado em relação a quem ouve a narrativa.

Com essa restrição, entretanto, pode-se dizer que o tempo semântico estabelece uma relação com o tempo cronológico tal como entendido extralingüisticamente: ou o tempo real, ou o tempo convencionado em uma narração. Desse ponto de vista, trata-se de uma categoria dêitica — ou seja, a referência do tempo de uma frase depende da situação em que a frase é enunciada. Assim, *chegarei de avião* remete a um tempo posterior àquele em que a frase é enunciada.

Em português, distinguem-se três tempos semânticos básicos, a saber, **presente**, **passado** e **futuro**. Os três têm representação formal, embora raramente exclusiva; por exemplo,

o **presente do indicativo** pode exprimir presente, passado ou futuro;

o **pretérito perfeito do indicativo** pode exprimir passado ou futuro;

o **pretérito imperfeito do indicativo** pode exprimir passado;

o **futuro do indicativo** pode exprimir futuro.

Além desses três tempos básicos, temos que considerar a referência não-temporal, exemplificada na frase

(21) A água ferve a 100 graus.

Aqui não se trata semanticamente nem de presente, nem de passado, nem de futuro, mas de uma afirmação geral que pretende valer para qualquer tempo. Tais frases ocorrem normalmente no presente do indicativo; em orações subordinadas, as exigências da correlação de tempos verbais podem fazer com que ocorram no pretérito imperfeito do indicativo, como em

(26) Manuel esqueceu que a água fervia a 100 graus.

Note-se que em (26) podemos substituir *fervia* por *ferve* sem mudança de significado.

Além de seu caráter dêitico (referência, mais ou menos direta, ao tempo cronológico), o tempo de um verbo em português pode ser definido em relação ao tempo de outro verbo da mesma oração, ou de outra oração do discurso. Assim, podemos ter

(27) Manuel tinha almoçado quando eu cheguei.

A referência de *tinha almoçado* se define em relação à de *cheguei*; isto é, ambos os fatos se deram no passado, mas o de almoçar se deu antes do de chegar. Como sabemos, há algumas formas especializadas na expressão de tempo anterior a outro do mesmo discurso; um exemplo é *tinha almoçado*, como em (27); outro é o pretérito mais que perfeito simples (*almoçara*).

A categoria semântica de "tempo" é altamente codificada em português; isto é, a língua possui formas mais ou menos especializadas para exprimi-la (a saber, os tempos verbais, assim como certas construções de auxiliar mais verbo principal). Mesmo quando o tempo semântico é claramente denotado através de elementos léxicos, como os chamados "advérbios de tempo", a codificação no verbo continua sendo exigida. Assim, podemos dizer

(28) Manuel picou a couve ontem de noite.

A presença do elemento *ontem de noite* deixa bem claro que o tempo semântico é passado em relação ao momento da enunciação. No entanto, isso não dispensa o verbo de ocorrer no passado:

(29) * Manuel pica/picará a couve ontem de noite.

Como veremos, não acontece o mesmo com as outras duas categorias que estudaremos abaixo (o aspecto e o modo).

9.4.3.2. Aspecto

É preciso diferenciar a noção de tempo semântico da de aspecto. Vejamos o seguinte par de frases:

(30) Meu tio escreveu um livro.

(31) Meu tio estava escrevendo um livro.

Há uma diferença semântica muito nítida entre elas, e não se trata de uma diferença de referência temporal, já que ambas se situam, semanticamente, no passado.

Trata-se de uma diferença de **aspecto**. Para diferenciar "aspecto" de "tempo semântico", vou utilizar a explicação dada por Comrie (1976), adaptando seus exemplos para o português. Diz Comrie que

> Os aspectos são maneiras diferentes de encarar a constituição interna de uma situação. [1976, p. 3]

Vejamos então a frase

(32) Cecília estava lendo quando eu entrei.

O fato veiculado por *entrei* é considerado globalmente:

> [a situação] é apresentada como um todo único, não analisável, com começo, meio e fim englobados em uma coisa só; não há tentativa de dividir a situação nas várias fases individuais que compõem a ação de entrar. [Comrie, 1976, p. 3]

Em contraste, *estava lendo* é considerado em seu desenvolvimento; em (32), *estava lendo* focaliza o meio da ação, e não o seu início ou fim. O resultado é que a interpretação semântica de (32) inclui a idéia de que eu "entrei" durante o desenvolvimento da ação de "ler", praticada por Cecília.

Como se vê, a diferença semântica entre as duas formas verbais não pode ser considerada uma diferença de tempo semântico. Dizemos que *escreveu* em (30), assim como *entrei* em (32), expressam o aspecto **perfectivo**; e *estava escrevendo* em (31), assim como *estava lendo* em (32), expressam o aspecto **imperfectivo**.

Aqui não será possível discutir em detalhe a conceituação de cada um dos aspectos encontrados em português ou nas línguas do mundo. Teremos de contentar-nos com alguns exemplos.

Pode-se entender o aspecto perfectivo como a expressão de um fato globalmente considerado, sem análise de suas fases, nem ênfase sobre uma ou outra dessas fases; já o imperfectivo inclui consideração das diversas fases, e por isso compreende várias modalidades. Assim, em

(30) Meu tio escreveu um livro.

temos a ação pura e simples, sem análise da situação em seus constituintes; o aspecto, portanto, é perfectivo. Já em

(31) Meu tio estava escrevendo um livro.

entende-se a ação no meio de seu desenrolar; o aspecto é imperfectivo — uma modalidade de imperfectivo geralmente chamada **progressivo**. Na frase

(33) Meu tio escrevia livros.

temos igualmente o imperfectivo, pois o fato de escrever livros é considerado em sua repetição: meu tio escreve livros com certa freqüência. Diremos então que a modalidade é **habitual**. Note-se a nítida diferença semântica entre (31) e (33), no que diz respeito à maneira como se encara a ação verbal.

9.4.3.3. Modo

Tradicionalmente se considera que os verbos também codificariam uma oposição semântica de **modo**; o modo se definiria semanticamente como caracterizando a "atitude do falante frente àquilo que está dizendo" — uma noção próxima, portanto, à de força ilocucionária, vista no capítulo 3.

Assim, teríamos basicamente três modos, a saber, o "indicativo", que exprimiria uma atitude de certeza do falante quanto ao que declara; o "imperativo", que veicularia ordens ou pedidos; e o "subjuntivo", que exprimiria uma atitude de incerteza, dúvida ou desejo frente ao conteúdo do enunciado. Exemplos seriam as ocorrências do verbo *lavar* em

(34) Selma afirmou que lavou o carro. (Indicativo)

(35) Lave o carro! (Imperativo)

(36) Quero que Selma lave o carro. (Subjuntivo)

(37) Duvido que Selma lave o carro. (Subjuntivo)

A julgar por esses quatro exemplos, a correlação entre a oposição morfológica usualmente denominada de "modo" (*lavou* X *lave* etc.) e uma oposição semântica como a que se delineou acima realmente funcionaria. Mas na verdade a correlação é tão frágil que exemplos como os que vimos constituem antes a exceção que a regra.

O fato, tal como o concebo, é que a oposição de modo (em especial a oposição indicativo/subjuntivo) tende, em português, a se tornar puramente formal. Na maioria dos casos, a oposição morfológica entre indicativo e subjuntivo é governada por traços semanticamente não motivados dos verbos (e de alguns outros itens, como *talvez*); os casos em que se pode ver um efeito semântico imputável ao modo são excepcionais e tendem a desaparecer na língua moderna.

Deixando de lado por ora o imperativo, vamos examinar brevemente as condições semânticas de ocorrência de subjuntivos e indicativos. As condições formais foram estudadas na seção 6.4.

O subjuntivo aparece principalmente em orações subordinadas. A oposição entre "certeza" e "incerteza", no entanto, não parece desempenhar um papel fundamental na determinação da ocorrência de cada um dos modos verbais. Por exemplo, temos

(38) Desconfio que Selma fuma cachimbo.

(39) Admito que Selma fume cachimbo.

Ambas as frases expressam uma certeza condicionada; não me parece que a segunda expresse mais incerteza que a primeira, e no entanto uma tem indicativo e a outra subjuntivo.

Nos exemplos

(40) Tenho certeza que Selma fuma cachimbo.

(41) É trágico que Selma fume cachimbo.

a posição do falante é de certeza quanto ao fato de Selma fumar cachimbo. Na primeira frase, essa certeza é afirmada, e na segunda é pressuposta, mas em ambas está presente. Mas, novamente, uma tem subordinada no indicativo e a outra tem subordinada no subjuntivo.

Finalmente, qualquer dos dois modos pode expressar uma atitude de falta de certeza:

(42) Eu sonhei que Selma fumava cachimbo.

(43) Eu duvido que Selma fume cachimbo.

Por outro lado, há pelo menos vestígios na língua de uma oposição semântica que, se não se identifica simplesmente com "certeza" *versus* "incerteza", é algo semelhante. Primeiro, a maioria dos verbos que exprimem dúvida realmente exigem subjuntivo. E há ainda alguns que exprimem certeza e exigem indicativo, mas que, quando negados, podem aceitar subjuntivo:

(44) Eu creio que Selma fuma cachimbo.

(45) a. Eu não creio que Selma fuma cachimbo.

b. Eu não creio que Selma fume cachimbo.

Outro vestígio de codificação da oposição "certeza"/"incerteza" está nas orações relativas ("adjetivas"). Nestas, a oposição subjuntivo/indicativo pode servir para diferenciar elementos tomados referencialmente (e, portanto, em geral admitidos como existentes) de elementos tomados atributivamente (admitidos como possíveis, mas não necessariamente existentes). Por exemplo,

(46) a. Procuro uma mulher que fuma cachimbo.

b. Procuro uma mulher que fume cachimbo.

O falante de (46a) refere-se a uma mulher tomada como existente e que está sendo procurada; já o de (46b) deseja encontrar uma mulher que corresponda à descrição dada, mas não se compromete quanto à sua existência.

A codificação exemplificada em (46a) e (b), porém, representa um uso pouco freqüente, e é possível que para muitos falantes as duas frases se-

jam, hoje, praticamente sinônimas. É mais um exemplo da tendência que se observa na língua a eliminar o papel semântico do subjuntivo (e, portanto, como sabemos, a eliminar gradativamente o próprio subjuntivo).

Voltando aos exemplos (34), (36) e (37), é mais indicado atribuir as diferenças de "certeza" à própria semântica do verbo principal (*afirmar, querer, duvidar*) do que ao modo do verbo subordinado.

Consideremos agora o chamado modo "imperativo". Este, com efeito, é bastante especializado na expressão dos valores ilocucionários de "ordem" ou "pedido". Não se pode utilizar um imperativo com outro valor; mas, como se viu na seção 3.1., ordens e pedidos podem perfeitamente ser veiculados por orações no indicativo:

(47) Você poderia fechar essa janela.

Aqui, portanto, vemos uma correlação, embora apenas unidirecional, entre modo verbal e modo semântico: o imperativo expressa a atitude de dar uma ordem ou fazer um pedido; entretanto, não é só o imperativo que tem essa função.

Finalmente, vale mencionar o caso das frases optativas, ou seja, as que têm verbo no subjuntivo mas não são subordinadas. Estas expressam sempre um desejo:

(48) Deus te ajude.

No entanto, são algo raras, e tendem a limitar-se a certas fórmulas feitas. E, naturalmente, um desejo pode ser (e geralmente é) expresso por frases não-optativas:

(49) Eu gostaria que Deus te ajudasse.

(50) Deus vai te ajudar, tenho certeza.

10 Papéis semânticos e funções sintáticas

Nesta seção vou abordar uma questão importante para a descrição semântica da língua: Como se interpretam semanticamente as diferentes funções sintáticas? Ou seja, o que resulta, no plano do significado, do fato de um sintagma ser sujeito, atributo ou objeto direto?

Algumas afirmações tradicionais, do tipo "objeto direto é o termo que exprime o ser que sofre a ação", parecem refletir a crença de que a resposta seria relativamente simples, com uma espécie de tabela dando o significado típico de cada função sintática. Na verdade, os fatos são vastamente mais complicados do que dá a entender essa posição tradicional. Aqui nos defrontamos mais uma vez com o problema que aponto repetidamente no decorrer deste trabalho: o caráter complexo e indireto da relação entre forma e significado.

Está por ser elaborada ainda uma análise de conjunto da relação entre as funções sintáticas e seu significado. A seguir, vou apenas expor e discutir alguns pontos selecionados, que deverão ser suficientes para que o leitor perceba a natureza e a complexidade da relação — e, espero, também para que ele se interesse pelo problema.

10.1. PAPÉIS SEMÂNTICOS

Vejamos primeiro uma breve conceituação de "papel semântico". Esse termo se refere às relações de significado expressas pelas funções sintáticas em si, independentemente de seu preenchimento léxico. Assim, pode-se falar do significado (ou melhor, dos significados) da função de sujeito; a idéia é que qualquer elemento que seja sujeito em determinadas condições terá uma relação semântica particular com o significado do restante da oração. O significado de *Zé quebrou a garrafa* não é apenas a soma dos significados dos diferentes constituintes *Zé* + *quebrou* + *a garrafa*; é necessário acrescentar que Zé praticou, e a garrafa sofreu, a ação expressa pelo verbo quebrou. Diremos

que o papel semântico de Zé nessa frase é de **agente**, e o da garrafa de **paciente**; ou seja, o sujeito dessa frase tem o papel semântico de agente, e o objeto direto o de paciente.

Agente é o papel semântico desempenhado pela entidade que provoca a ação denotada pelo verbo; entende-se que o agente provoca essa ação voluntariamente. Algumas conseqüências dessa definição são que só pode ocorrer agente quando a sentença exprime uma ação; e que o agente é sempre um ser animado (humano ou animal).

Paciente é o papel semântico que expressa a entidade diretamente afetada por uma ação, ou que "sofre" a ação, para utilizar o termo tradicional.

Instrumento é o papel desempenhado pela entidade (em geral, um objeto inanimado) utilizada pelo agente para levar a efeito sua ação.

Fonte é o papel que expressa a origem de um movimento, como *de São Paulo* em *Marcinha dirigiu de São Paulo até Arapongas*; mas também se aplica ao movimento figurado, como *Marcinha* em *Marcinha vendeu a charrete para um alemão*: a charrete se "movimentou" da posse de Marcinha para a posse do alemão.

Meta expressa o final de um movimento real ou figurado; assim, nas duas frases do parágrafo precedente, *até Arapongas* e *para um alemão* têm o papel de meta.

Não são somente esses os papéis semânticos já distinguidos pelos lingüistas; naturalmente, é de se esperar que haja muitos, pois as noções a serem caracterizadas também são numerosas. No entanto, neste capítulo, só lidaremos com os cinco papéis dados acima.

É preciso notar, finalmente, que um único sintagma pode receber simultaneamente mais de um papel semântico; assim, em *Marcinha vendeu a charrete para um alemão*, o SN *Marcinha* é ao mesmo tempo agente da venda e fonte da qual a charrete se transferiu para a posse do alemão.

O leitor provavelmente notará que as definições dos papéis semânticos são vagas; na verdade, uma das tarefas (e um dos desafios) que a lingüística precisa enfrentar urgentemente é a de conceituar os papéis semânticos com mais rigor. Por enquanto, é inevitável que haja certo grau de vaguidão nas definições.

10.2. SEMÂNTICA DE TRÊS VERBOS

Vamos tomar como exemplos os verbos *quebrar*, *matar* e *comer*, cujo comportamento semântico é parcialmente diferente. Tentaremos construir um mecanismo capaz de interpretar semanticamente as diferentes funções sintáticas nas construções em que entram esses três verbos. As conclusões devem valer igualmente para outros verbos (e o leitor é convidado a investigar quais); mas não

haja dúvida de que a análise está muito longe de dar conta de todos os casos de frases com verbo de ação. Há pouca pesquisa relevante neste setor; as páginas seguintes devem servir de ilustração do trabalho a fazer e de sua complexidade.

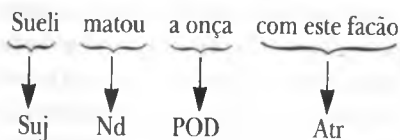
10.2.1. Semelhanças entre os verbos

Observemos inicialmente que a transitividade dos três verbos é a mesma: são todos os três marcados como [L-OD, L-AC, Rec-Pv, Rec-CP] (ver o capítulo 6 para a conceituação de "transitividade"). E comecemos com os seguintes exemplos:

(1) Sueli quebrou a vidraça com este martelo.

(2) Sueli matou a onça com este facão.

(3) Sueli comeu a pizza com este garfo.

Analisando essas três orações em termos das funções do capítulo 3, verificamos que são sintaticamente paralelas, contendo cada uma sujeito, núcleo do predicado, objeto direto e atributo, como no exemplo

A interpretação semântica também é paralela, pois nas três frases vale a seguinte tabela de relação entre funções sintáticas e papéis semânticos:

Suj: agente;
OD: paciente;
Atr: instrumento.

Até o momento, portanto, as coisas parecem ser bastante simples. Se omitimos o atributo, deixa de ser expresso o instrumento, mas as relações permanecem inalteradas:

(4) Sueli quebrou a vidraça.

(5) Sueli matou a onça.

(6) Sueli comeu a pizza.

O sujeito continua sendo agente, e o objeto direto o paciente.

Agora, nas frases seguintes a coisa se complica:

(7) Este martelo quebrou a vidraça.

(8) Este facão matou a onça.

Embora essas frases sejam perfeitamente aceitáveis, aqui o sujeito é interpretado como **instrumento**. O agente fica sem especificação e se entenderá geralmente como desconhecido ou irrelevante para a mensagem. Já o objeto direto continua sendo interpretado como paciente.

Como é evidente, a tabela precisa ser modificada para dar conta dos novos fatos. Digamos que a modificação se faça da seguinte forma: o sujeito pode ser interpretado (mas só em frases sem atributo) como

agente ou como instrumento, indiferentemente. Vimos que em

(5) Sueli matou a onça.

o sujeito é agente; e em

(8) Este facão matou a onça.

o sujeito é instrumento.

Pela modificação introduzida, seria de esperar que cada uma dessas frases fosse ambígua, com o sujeito sendo agente ou instrumento em cada uma. Mas a ambigüidade potencial é eliminada em virtude de condições de boa formação semântica: em (5), o sujeito é interpretado como agente e não como instrumento, porque é *Sueli*, que se refere a uma pessoa; e em geral não estamos preparados para ver uma pessoa utilizada como objeto para matar uma onça ou quebrar uma vidraça. Já em (8) o sujeito é *este facão*, que se refere a um objeto inanimado e não poderia ser agente; conseqüentemente, entende-se como instrumento.

Essa análise prevê que em circunstâncias em que as restrições semânticas mencionadas forem relaxadas a interpretação possível do sujeito dessas frases poderá mudar. Por exemplo, se (5) for inserida em um contexto em que Sueli, desmaiada, cai em cima da onça, matando-a, poderemos entender Sueli como instrumento; ou então se (8) for colocada no contexto de uma fábula, onde o facão se apresenta como personagem ativa, poderemos entendê-lo como agente.

10.2.2. Diferenças

Por outro lado, a interpretação "agente ou instrumento" não vale para o sujeito em geral. Primeiro, como já vimos, o sujeito só pode ser instrumento se não houver um atributo que ocupe esse papel: é o caso das frases (1) a (3), onde, em qualquer circunstância, a interpretação de Sueli é necessariamente como agente. Além disso, há frases como as seguintes:

(9) A vidraça quebrou.

(10) Onça mata.

(11) Sueli comeu.

Até agora, pelo menos *quebrar* e *matar* se vinham comportando da mesma maneira: o que valia para frases com um desses verbos valia igualmente para frases com o outro. Mas aqui o paralelismo semântico já não se verifica. Em (10) e (11), o sujeito é agente, como seria de esperar, mas em (9) o sujeito é paciente: a vidraça sofre a ação de quebrar.

Vou propor uma análise logo adiante para dar conta desses fatos e da interpretação geral de frases com esses três verbos. Conforme veremos, a análise evidencia também outros fatos interessantes sobre a semântica desses verbos e servirá para ilustrar o funcionamento do componente semântico da gramática e do léxico. A exposição da análise e sua aplicação a exemplos é algo comple-

xa, o que é inevitável, dada a complexidade dos próprios fatos.

Comecemos depreendendo alguns traços semânticos dos verbos em questão. Primeiro, é fácil mostrar que *quebrar* exige a presença do paciente em sua representação semântica; assim, frases como as seguintes são inaceitáveis:

(12) * Sueli quebrou.

Essa frase só se torna aceitável se imaginarmos Sueli como uma boneca de louça, por exemplo; nesse caso seria paciente, e não haveria razão para sentir a frase como mal formada (não nos interessa aqui a acepção de *quebrar* como "ir à falência"). Mas, fora dessa interpretação um tanto forçada, a representação semântica de (12) carece de paciente. E, como isso resulta em inaceitabilidade, podemos afirmar que

[*quebrar* exige paciente].

Essa exigência é um dos muitos traços semânticos que caracterizam o significado do verbo *quebrar*.

Passando agora ao verbo *matar*, é possível mostrar que ele exige um agente ou um instrumento. Assim, podemos dispensar a expressão do paciente e do instrumento, como em

(10) Onça mata.

Aqui o agente foi expresso. Podemos exprimir o instrumento, dispensando o paciente e o agente, como em

(13) Este facão mata.

Mas não há maneira de exprimir apenas o paciente, ou nenhum dos três papéis semânticos. Ou seja, não há correspondente com *matar* de uma frase como (9), onde apenas o paciente é expresso:

(9) A vidraça quebrou.

Assim, podemos atribuir a esse verbo o traço semântico:

[matar exige agente ou instrumento].

Pode-se perguntar: e se alguém quiser exprimir um ato de matar sem explicitar o agente nem o instrumento? Será impossível? A resposta é que a língua oferece outros recursos para isso: pode-se utilizar uma construção passiva, que tem como uma de suas funções justamente exprimir privilegiadamente o paciente: *a onça foi morta*; ou então pode-se utilizar outro verbo, que admita a expressão do paciente sozinho: *a onça morreu*.

Finalmente, o verbo *comer* também tem um traço semântico que nos será útil:

[*comer* exige agente].

(Veremos exemplificação disso mais adiante.)

Esses traços são parte integrante da semântica dos itens léxicos *quebrar*, *matar* e *comer*. Devem ser entendidos como condições de boa

formação das representações semânticas em que figuram os verbos respectivos. Assim, por exemplo, se a representação semântica de uma frase inclui *quebrar* (ou, mais exatamente, sua matriz semântica) e não inclui um paciente, então essa representação semântica é mal formada, e a sentença correspondente será inaceitável.

Os três verbos que estamos examinando têm traços semânticos distintos. Esses traços são responsáveis por diferenças de aceitabilidade de muitas frases que, em termos de sintaxe, parecem inteiramente paralelas. Passemos ao exame de um desses casos.

Vamos investigar a razão da estranheza de

(14) * Este garfo comeu a pizza.

Essa frase parece muito semelhante a

(8) Este facão matou a onça.

Examinando (8), vemos que o sujeito é instrumento, o objeto direto é paciente e o agente não foi expresso. A frase é perfeitamente aceitável. Ora, por que não aconteceria o mesmo com (14)? Teríamos como instrumento *este garfo* (que é semanticamente compatível com o papel de instrumento) e como paciente *a pizza*. À primeira vista, nada há que desautorize (14).

No entanto, há uma diferença semântica entre os dois verbos: *matar*, como vimos, exige a expressão do agente ou do instrumento, ao passo que *comer* exige a expressão do agente. Acontece que a interpretação semântica de (14) não inclui um agente, porque este teria de ser o sujeito, e o sujeito denota um ser inanimado, portanto inadequado ao papel de agente. A interpretação de (8) igualmente não tem agente, mas o verbo não o exige: *matar* exige a expressão de agente ou instrumento. Acontece que *este facão* em (8) pode ser interpretado como instrumento, e dessa forma a exigência do verbo fica satisfeita.

10.2.3. Regras semânticas

Eu disse acima, falando da frase (14), que o agente "teria de ser o sujeito". Mas como sabemos isso? Acontece que existem regras que estabelecem esse tipo de relação: são chamadas **regras de interpretação semântica**, ou mais simplesmente **regras semânticas**. Vou agora apresentar um pequeno conjunto dessas regras, suficientes para descrever os fenômenos ligados às frases com os três verbos que estamos considerando (mas não suficientes para descrever as relações entre funções sintáticas e papéis semânticos em geral!).

As regras que utilizaremos são três, a saber:

Regra 1 (R1): O objeto direto se interpreta como paciente.

Regra 2 (R2): O atributo (com a preposição *com*) se interpreta como instrumento.

Regra 3 (R3): O sujeito se interpreta como: agente, instrumento ou paciente (nessa ordem de preferências).

Essas três regras devem ser aplicadas na ordem dada: primeiro a Regra 1, depois a 2, depois a 3.

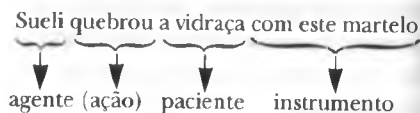
Além disso, convém lembrar que o agente precisa denotar um ente animado.

Vejamos agora alguns exemplos de interpretação de frases. Como se verá, a interpretação depende da aplicação das regras semânticas; dos traços semânticos dos verbos (e dos outros elementos da frase); e de certos princípios gerais, que serão mencionados à medida que for necessário.

Seja a frase

(1) Sueli quebrou a vidraça com este martelo.

Primeiro aplicamos as três regras; a Regra 1 (R1) marca o OD *a vidraça* como paciente; R2 marca o atributo *com este martelo* como instrumento; e R3 marcará o sujeito *Sueli* como agente — não só porque este é o primeiro papel na ordem de preferências, mas também porque os outros dois papéis já estão ocupados. Um princípio que podemos adotar desde já é o de que cada papel semântico (dentre os três estudados) só aparece uma vez em cada oração. Assim, o resultado até o momento é

Sueli quebrou a vidraça com este martelo
↓ agente | ↓ (ação) | ↓ paciente | ↓ instrumento

Ainda é preciso verificar se as exigências semânticas do verbo estão satisfeitas. *Quebrar*, como sabemos, exige paciente; na interpretação acima, o paciente está preenchido pelo constituinte *a vidraça* (ou, mais exatamente, pela interpretação semântica de *a vidraça*). Portanto, não há violação por este lado.

Outra coisa que é preciso verificar é se o significado dos diversos elementos está de acordo com o papel que cada um desempenha. Por exemplo, o agente precisa ser animado; no caso, é *Sueli*, nome de pessoa, e portanto tampouco aqui há violação. Como resultado, a interpretação é bem formada, e a frase é aceitável tanto sintática quanto semanticamente.

Passemos agora a outro exemplo:

(9) A vidraça quebrou.

R1 não pode ser aplicada, porque não há objeto direto na oração; o mesmo acontece com R2, porque não há atributo. Passamos à aplicação de R3.

R3 marcaria o sujeito como agente; mas o resultado seria semanticamente malformado, porque o agente precisa ser animado, e *a vidraça* denota ser inanimado. Portanto, essa interpretação é automaticamente excluída (ou "filtrada"), e passamos à segunda: o sujeito seria instrumento. *A vidraça* poderia ser instrumento, mas acontece que *quebrar* exige paciente especificado. Se *a vidraça* for instrumento, não restará nenhum elemento que possa ser interpretado como paciente; a representação semântica final será, portanto, mal formada, por conter *quebrar* e não conter um paciente especificado. Assim, ainda essa alternativa acaba sendo filtrada.

Finalmente, R3 interpreta *a vidraça* como paciente. Nesse caso, as exigências de *quebrar* ficam satisfeitas, a interpretação é bem formada e a frase é aceitável.

Agora o sistema terá de dar conta da inaceitabilidade de

(15) * A pizza comeu.

R1 e R2 não se aplicam, por falta de objeto direto e de atributo. Passamos a R3: a primeira alternativa é interpretar o sujeito como agente, mas esta é excluída porque *a pizza* é inanimado e não pode ser agente. A segunda possibilidade é que *a pizza* seja instrumento, mas já sabemos que *comer* exige agente, e se *a pizza* for instrumento não poderá haver especificação do papel de agente, pois não sobrará nenhum elemento para assumi-lo. O resultado será mal formado, e portanto será filtrado. Finalmente, R3 pode interpretar *a pizza* como paciente: mas, ainda assim, o agente ficaria sem representação semântica, desobedecendo a exigência de *comer*. Terminadas as alternativas, vemos que nenhuma delas dá resultado semanticamente bem formado; a interpretação semântica, qualquer que seja, é anômala e a frase é inaceitável.

Creio que com esses exemplos fica bastante claro como se aplica o mecanismo da interpretação semântica. Esse mecanismo atribui significados a estruturas bem formadas e filtra aquelas cuja interpretação apresenta alguma anomalia. O leitor poderá aplicar as regras e traços semânticos a outros exemplos, como os seguintes:

(2) Sueli matou a onça com este facão.

(6) Sueli comeu a pizza.

(12) * Sueli quebrou.

(14) * Este garfo comeu a pizza.

(16) Sueli já comeu.

etc.

Em todos eles, as regras deverão produzir a interpretação correta para cada um dos termos estudados (sujeito, objeto direto e atributo); e nos casos de inaceitabilidade o mecanismo deverá excluir todas as alternativas possíveis, ao mesmo tempo diagnosticando a causa ou as causas da inaceitabilidade.

10.2.4. Que é uma sentença aceitável?

Já sabemos que uma sentença aceitável é aquela que é bem formada segundo os diversos componentes da gramática: fonológico, morfológico, sintático e semântico. Assim, uma frase morfologicamente mal formada é a que inclui palavras como * *comimos*, que viola as regras da morfologia; uma frase sintaticamente mal formada viola as regras da sintaxe, como, por exemplo,

(17) * Eu gosto vocês todos.

Aqui não há problema de significado, mas a regência verbal está errada.

Finalmente, uma frase pode ser semanticamente mal formada, por conter em sua interpretação semântica alguma anomalia, como

(18) * O facão comeu a pizza.

(Onde, exatamente, está a má formação dessa frase?)

10.2.5. Interpretação das estruturas passivas

As regras 1 a 3, vistas na seção 10.2.3., descrevem a interpretação das funções sintáticas de muitos verbos, na maioria das construções em que estes entram. No entanto, há um grupo de construções que requerem tratamento especial: as orações passivas. Nesta seção, vamos ver como se interpretam essas construções.

O problema principal tem a ver com a introdução do agente na representação semântica das sentenças. Se submetermos uma frase como

(19) Meu livro foi rasgado.

à aplicação das regras 1 a 3, deveremos atribuir ao sujeito, *o livro*, o papel semântico de agente; evidentemente, a interpretação resultante não estaria de acordo com o julgamento dos falantes, para quem o sujeito é claramente paciente.

Se, de alguma maneira, marcarmos o sujeito de (19) como paciente, ainda assim restará um problema: a Regra 3 atribuirá ao sujeito (*meu livro*) o papel de instrumento (que tem preferência sobre o de paciente), resultando em outra interpretação errônea.

A solução que vejo no momento é construir um sistema de interpretação especial para as frases passivas. Idealmente, esse sistema deveria ser integrado ao sistema constituído pelas regras 1 a 3; mas, como não se sabe a maneira de fazer isso, aqui nos contentaremos com dois sistemas paralelos, que se aplicam a dois grupos separados de estruturas.

Vamos admitir, então, que o mecanismo de interpretação começa inspecionando a estrutura de cada sentença, para averiguar qual dos dois sistemas deverá ser encarregado de interpretá-la. Isso feito, a sentença será submetida às regras 1 a 3 ou,

então, caso sua estrutura seja do tipo passivo, às novas regras que formam o sistema paralelo.

Começamos, portanto, formulando o mecanismo de inspeção das estruturas. Os dois tipos se distinguem através da aplicação do seguinte mecanismo de rotulação:

Definição das estruturas passivas

— *Se a estrutura contém predicado composto do verbo ser + particípio, é uma estrutura* passiva.

— *Caso contrário, é uma estrutura* ativa.

Agora cada tipo de estrutura dará entrada no seu conjunto específico de regras de interpretação. As estruturas ativas são interpretadas pelas regras 1 a 3, tais como definidas e explicadas na seção 10.2.3. Já para as estruturas passivas será preciso formular novas regras, a saber:

Regras para a interpretação de estruturas passivas (regras P)

Regra 1-P (R1-P): O adjunto circunstancial (com *por*) se interpreta como agente.

Regra 2-P (R2-P): Introduzir "agente não-especificado" na representação semântica das passivas, caso essa representação não inclua "agente".

Regra 3-P (R3-P): O sujeito se interpreta como paciente.

Essas regras se aplicam na ordem dada.

Vejamos agora alguns exemplos de aplicação desse mecanismo. Seja a frase

(20) Meu livro foi rasgado.

A definição rotula a estrutura como passiva, porque seu predicado é composto de *ser* + particípio (este por sua vez se define morfologicamente, pela presença do sufixo característico). Assim, a interpretação se fará pelo sistema das Regras P.

A Regra 1-P não se aplica, porque não há adjunto circunstancial com *por.* Se a frase fosse *meu livro foi rasgado pelos garotos,* naturalmente *pelos garotos* receberia o papel de agente.

Aplicamos agora R2-P: o que temos em mão no momento é uma frase passiva e sua representação semântica não inclui agente. Assim, atribuímos a essa representação um agente não-especificado.

Finalmente, R3-P interpretará o sujeito, *meu livro,* como paciente. O resultado é uma representação semântica de acordo com a intuição dos falantes: *meu livro* é o paciente e o agente é não-especificado.

Como se vê, as passivas precisam ser interpretadas por um sistema separado; isso é, repito, provavelmente, uma contingência do precário estado de nosso conhecimento do assunto. É necessário procurar meios de integrar a interpretação das passivas no sistema geral de interpretação (talvez lançando mão da relação de correspondência parcial existente entre passivas e ativas;

ver a seção 7.3.1. a respeito dessa relação). No momento, temos de conformar-nos com essa solução provisória, que pelo menos fornece uma descrição razoavelmente adequada dos fatos. E que esta (auto)crítica fique como sugestão de pesquisa.

10.2.6. Impessoais

Examinemos agora o caso de frases "impessoais", das quais distingo dois tipos, exemplificados pelas frases seguintes:

(21) Quebraram meu relógio.

(22) Come-se pizza no Natal.

Vamos considerar primeiro o exemplo (21). Como não se trata de uma passiva, essa frase deverá ser interpretada pelo primeiro conjunto de regras, R1-R3. Assim, aplicando R1, o objeto direto *meu relógio* recebe o papel semântico de paciente. R2 não se aplica, pois não há atributo.

Mas ao chegarmos a R3 surge um problema: a frase não tem sujeito, logo R3 não se aplica. Como resultado, a representação semântica de (21) não terá agente — mas isso vai contra nossa intuição. Entendemos (21) como necessariamente tendo agente; por exemplo, não se pode interpretar (21) como significando que o relógio se quebrou por ter sido atirado do quarto andar pelo vento.

Aqui, portanto, encontramos uma falha do sistema desenvolvido até o momento. Vamos saná-la introduzindo uma outra regra.

Regra 4 (R4)

Introduzir "agente (não-especificado)" na representação semântica de frases com NdP na 3ª pessoa do plural, sem sujeito.

(Por razões que veremos na seção 11.2.2., esta regra se define como de aplicação **opcional**.)

O leitor deverá ter notado que esta já é a segunda regra que introduz "agente não-especificado" em uma representação semântica; a outra é R2-P, que introduz o agente não-especificado em casos de passivas sem agente. Conforme veremos, há ainda pelo menos uma outra estrutura que recebe essa interpretação para o agente (ver a seção 11.2.4.4.). Não vou tentar aqui unificar os três processos, mas talvez isso seja possível, pois as semelhanças entre eles são suspeitas.

Voltando ao exemplo (21),

(21) Quebraram meu relógio.

já vimos como R1 atribui ao OD o papel de paciente e como R2 deixa de se aplicar por falta de atributo. R3 também não se aplica, porque não há sujeito. Agora, então, aplicamos R4; a estrutura está de acordo com a descrição estrutural fornecida na regra: (21) tem o NdP na terceira pessoa do plural e não tem sujeito. Por conse-

guinte, introduzimos na representação semântica um agente não-especificado, obtendo a acepção correta, algo como "uma entidade animada não-identificada quebrou meu relógio". (Para ser mais exatos, deveríamos ter especificado que o agente, além de animado, é humano.)

Restaria agora discutir a maneira de interpretar:

(22) Come-se pizza no Natal.

Esse tipo de oração apresenta a dificuldade de não ter uma análise clara; em especial, não é claro qual é o sujeito: é *pizza*, ou a oração não tem sujeito?

Isso acontece porque há muita flutuação entre os falantes (e escritores) sobre se se deve dizer:

(23) Come-se pizzas no Natal.

(24) Comem-se pizzas no Natal.

Os gramáticos normativos aconselham (24), e muita gente os segue. Mas isso deve ser um fenômeno basicamente escolar; não creio que haja pessoas que aprendam nativamente a produzir frases como (24). Desse modo, uma gramática do português que consigne (24) como a única forma aceitável está escondendo fatos [e, também, uma gramática que condene (23) como incorreta está perdendo tempo]. Aqui, para contentar a todos, vou esboçar a maneira de interpretar tanto (23) quanto (24).

Tomemos primeiro (23). A maneira de incluir essa frase em nossa análise é estender R4, de maneira que introduza o agente especificado na representação semântica de frases com NdP na terceira pessoa mais *-se*. Nesse caso, a interpretação de (23) será: R1 atribui ao objeto direto (*pizzas*) o papel de paciente; R2 não se aplica, pois não há atributo; R3 tampouco, pois também não há sujeito. Finalmente, R4 introduz o papel de agente (não-especificado). O resultado, como o leitor facilmente pode verificar, é uma interpretação correta.

Agora vamos considerar o caso de (24). A diferença crucial aqui é que *pizzas*, que é o paciente, é sujeito (o verbo concorda com ele). Desse modo, a maneira mais fácil de interpretar (24) é fazer com que sofra a ação das regras P, especializadas na interpretação das passivas. Isso não significa que tenhamos que incluir (24) na definição de "passiva", o que poderia ser difícil, dada a grande diferença formal entre as duas construções. Mas podemos estender a ação das regras P de maneira a se aplicarem não apenas às passivas, mas ainda a construções do tipo de (24). Segundo essa análise, a semelhança entre (24) e as passivas — captada pela gramática tradicional quando chama (24) de "passiva sintética" — estaria principalmente no fato de que o sujeito se interpreta como paciente.

Nos termos da proposta acima, portanto, podemos aplicar a (24) as regras P. R1-P não se aplica, porque não há adjunto circunstancial; R2-P introduz um agente não-especificado; e R3-P atribui ao sujeito, *pizzas*, o papel de paciente. Também dessa maneira se obtém a interpretação correta.

Qual é a melhor análise, então? Depende de que estrutura se está interpretando. Se verificarmos que em português atual só existe (23), e não (24), a análise correta será a primeira; teremos, portanto, que estender R4 para introduzir agente não-especificado também em casos de 3ª do singular mais *-se*. Se apurarmos que só (24) é que existe, então teremos que estender a ação das regras P a essa construção. Finalmente, se descobrirmos que ambas as construções coexistem no português de hoje, teremos que fazer ambas as modificações: tanto em R4 quanto na ação das regras P. Aqui não vou tentar resolver essa questão; fica para pesquisa futura.

Uma advertência, porém (talvez desnecessária): não se trata de decidir qual das duas construções **deve** ser utilizada. Existe um uso estabelecido, que precisa ser investigado nos textos escritos em português padrão; essa investigação deve revelar qual é a forma ou formas correntes, que, portanto, merecem figurar nas gramáticas. Uma atitude normativa nesta questão, embora tradicional, seria totalmente equivocada.

10.2.7. Presença do agente nas representações semânticas

A análise acima proposta nos levou a lidar com agentes oriundos de duas fontes: em alguns casos, "agente" é o papel semântico atribuído a um sintagma explícito, formalmente presente na estrutura; em outros casos, o papel de agente foi introduzido em uma representação semântica sem se vincular diretamente a nenhum constituinte formal (embora, naturalmente, essa introdução sempre seja controlada por algum traço formal; por exemplo, pela presença de um NdP sem sujeito e na terceira pessoa do plural).

Por outro lado, em alguns casos, não há a introdução de nenhum agente na representação semântica. Assim, nossa análise atribui um agente não-especificado à representação semântica de

(20) Meu livro foi rasgado.

mas não atribui nenhum agente à representação semântica de

(9) A vidraça quebrou.

Naturalmente, é importante dar alguma razão para essa diferença de tratamento.

A primeira razão que se pode invocar é a intuição direta dos falantes. Parece-me razoavelmente claro que em (20) se subentende (ou, melhor, se pode subentender) um

agente por trás da ação descrita; pelo menos a interpretação mais imediata não é a de um rasgamento puramente acidental, digamos durante uma colisão de carros. Já em (9) a interpretação mais imediata é justamente a de um acontecimento cujo agente não existe, ou não tem importância para a mensagem. Para usar a terminologia vaga tantas vezes inevitável nos estudos semânticos, o significado de (20) focaliza de preferência uma ação, e o de (9) um simples acontecimento.

Concordo que essa argumentação repousa em bases pouco sólidas. Assim, vejamos outra razão, mais palpável, para distinguirmos (20) de (9) em termos da presença de um agente em suas interpretações semânticas. Existem certos elementos (que aparecem em geral na função de atributo) cujo significado é tal que só podem ocorrer se a frase incluir um agente em seu significado. Um desses elementos é a palavra *propositalmente* (outros exemplos são: *de propósito, para me incomodar, sem querer* etc.). Assim, uma frase como

(25) Marília fechou a janela.

que exprime uma ação, aceita o acréscimo desses elementos:

(26) Marília fechou a janela propositalmente.

Mas na frase

(27) Marília me parece muito gorda.

o acréscimo de *propositalmente* produz um efeito estranho:

(28) ?? Marília me parece muito gorda, propositalmente.

A estranheza vem de que estaríamos atribuindo um propósito a um fato que não depende da decisão de um agente (tanto "ser gorda" quanto "parecer-me gorda"). Portanto, podemos dizer que elementos como *propositalmente* só ocorrem sem estranheza em frases cuja interpretação inclua agente. Esses elementos são chamados **adverbiais orientados para o agente**.

Voltemos agora aos exemplos (20) e (9); utilizando o teste de *propositalmente*, fica bem evidenciada a diferença entre as duas frases; tudo se passa como se (20) tivesse um agente em sua interpretação, enquanto que (9) não o tem:

(29) Meu livro foi rasgado propositalmente.

(30) ?? A vidraça quebrou propositalmente.

(30) só poderia ter o estranho significado de que a vidraça decidiu, por sua própria volição, quebrar-se.

A análise desenvolvida nas seções precedentes para a interpretação das frases com *quebrar, matar* e *comer* explica esses fatos: (30) é estranha porque inclui um adverbial orientado para o agente, mas não inclui agente. (No Quadro 10-A, forneço mais alguma discussão sobre a aplicação desse teste.)

10.3. SUMÁRIO: PAPÉIS SEMÂNTICOS E FUNÇÕES SINTÁTICAS

Acabamos de ver um exemplo de como se descreve a relação existente entre as formas lingüísticas e seu significado. Vimos que as estruturas formais, produzidas com base nos componentes morfológico e sintático da gramática, sofrem a aplicação de regras semânticas, que lhes atribuem interpretações. Nesta seção, nosso exemplo foi a atribuição de papéis semânticos às funções sintáticas; nas seções seguintes deste capítulo, veremos outros exemplos da interpretação das estruturas da língua.

O exemplo dado foi extremamente restrito: restringimo-nos ao exame de três papéis semânticos em relação com três funções sintáticas; e limitamo-nos à interpretação das frases que contêm os verbos *quebrar, matar* e *comer.* O sistema se aplica a bom número de outros verbos, cujo comportamento é semelhante ao daqueles três; assim, *rasgar, entortar, abrir, fechar* e muitos outros verbos seguem o modelo de *quebrar.* No entanto, há também bom número de verbos para os quais o mecanismo proposto não funciona. Isso é um problema, claro, mas só poderá ser resolvido através da realização de novas pesquisas, dirigidas à obtenção de uma visão cada vez mais abrangente do fenômeno da interpretação das funções sintáticas e à elaboração de uma análise suficientemente integrada do fenômeno. Enquanto tal não se fizer, fique como exemplo a análise parcial dada nesta seção.

10.4. FUNÇÕES SINTÁTICAS E REFERÊNCIA

Agora vou discutir muito brevemente um outro aspecto da interpretação das funções sintáticas: as restrições que estas impõem ao tipo de referência que os sintagmas podem veicular.

Primeiro, observarei que um SN pode ser interpretado como referencial, atributivo ou qualificativo; nos exemplos abaixo, os sintagmas em itálico ilustram cada uma dessas acepções:

(31) *Um elefante* pisou no meu pé. [**Referencial**]

(32) Só *um elefante* daria conta de comer isso. [**Atributivo**]

(33) Dumbo é *um elefante.* [**Qualificativo**].

Entende-se que um sintagma tem interpretação **referencial** quando denota uma entidade entendida como existente, identificável; em (31), seria possível, por exemplo, identificar o elefante que pisou no meu pé. A interpretação **atributiva**, ao contrário, fornece apenas a descrição de uma entidade virtual, que

pode nem existir. Assim, em (32), *um elefante* está apenas descrevendo a entidade que conseguiria comer tudo: se for um elefante (qualquer que seja), satisfará essa condição. Note-se que (32) nem sequer exige que elefantes existam. Finalmente, a interpretação **qualificativa** como que seleciona apenas as qualidades relevantes da entidade mencionada: (33) não denota nenhum elefante identificável, nem sequer virtual; antes, atribui a Dumbo as qualidades (quaisquer que sejam) de um elefante.

O que nos interessa aqui é a seguinte observação: a função sintática de um sintagma pode condicionar o tipo de referência que ele pode ter. Em (33), o complemento do predicado tem interpretação qualificativa; o objeto direto pode ser atributivo, como em

(34) Eu gostaria de ter *um elefante.*

ou ainda referencial, como em

(35) Çãozinha comprou *um elefante.*

O sujeito, embora possa ser atributivo [como em (32)] ou referencial [como em (31)], nunca pode ser qualificativo. Se alterarmos a estrutura de (33), trocando de função os dois SNs, obteremos

(36) Um elefante é Dumbo.

No entanto, a interpretação qualificativa de *um elefante* fica perdida, justamente porque esse sintagma é agora o sujeito. Assim, enquanto (33) significa que Dumbo pertence ao conjunto dos elefantes (ou, talvez, tem as qualidades de um elefante), (36) só pode significar que um dentre vários elefantes mencionados é Dumbo.

Assim, verificamos que a função sintática é um dos fatores que condicionam o tipo de referência (há outros fatores, alguns de natureza extragramatical, que não precisamos considerar aqui). Um objeto direto ou um sujeito pode ser referencial ou atributivo. Um complemento do predicado pode ser qualificativo. Mas nem o sujeito nem o objeto direto podem ser qualificativos.

Podemos acrescentar que um predicativo só admite a interpretação qualificativa, nunca a referencial ou a atributiva. Por exemplo, em

(37) Elaine considera Ricardo um elefante.

não nos estamos referindo a um elefante pressuposto como existente, nem estamos falando de todo e qualquer elefante que possa existir. Estamos, em vez disso, afirmando que Elaine atribui a Ricardo as qualidades — ou os defeitos — de um elefante.

Outra função que parece só aceitar interpretação qualificativa é a de atributo. Em exemplos como

(38) Sônia comeu a pizza uma fera.

o sintagma *uma fera* é claramente qualificativo. Nisto (mas apenas nisto), a semântica do atributo é idêntica à do predicativo.

O assunto oferece, evidentemente, muitos temas de estudo. Como a pesquisa ainda não foi realizada, fiquemos por ora com a observação, já de si interessante, de que a função sintática é um dos fatores condicionantes do tipo de referência que um sintagma pode ter. Temos aqui mais uma faceta da correlação que existe entre os traços formais (funções sintáticas) e as propriedades semânticas (tipos de referência) das frases da língua.

Uma especulação interessante (mas no momento apenas uma especulação) seria sugerir que as restrições ao tipo de referência impostas pelas funções são responsáveis pelo fato de que algumas funções podem ser preenchidas por sintagmas nominais ou adjetivos (objeto direto) e outras apenas por sintagmas nominais (sujeito). Sabemos que o complemento do predicado pode ser nominal ou adjetival:

(39) Carlos é *o chefe da seção.*

(40) Carlos *está doente.*

Mas o sujeito só pode ser nominal, nunca adjetival. Isso poderia provir do fato de que sintagmas adjetivos não admitem interpretação referencial nem atributiva. Logo, um sintagma adjetivo que fosse sujeito não poderia ser completamente interpretado (faltaria o tipo de referência), e o resultado seria necessariamente inaceitável.

Como disse, trata-se apenas de uma especulação. Seria indispensável examinar um grande número de exemplos, assim como tentar generalizar a explicação para todas as funções e para outros traços semânticos. De qualquer modo, a hipótese de que muitos traços da sintaxe são, em última análise, o resultado formal de restrições semânticas é uma das mais interessantes direções de pesquisa que vejo atualmente na área. Depende de um desenvolvimento maior dos estudos semânticos e de uma integração cuidadosa destes com a pesquisa em sintaxe.

11 Elementos anafóricos

11.1. PRONOMES

Vamos chamar aqui **pronomes** os "pronomes pessoais" da gramática tradicional, ou seja, *eu, você, vocês, ele, ela, eles, elas, nós* etc., assim como os oblíquos e reflexivos *me, nos, lhe, o, os, se* etc. (conforme se verá no capítulo 13, os pronomes, assim definidos, são uma subclasse dos substantivos). Essas palavras têm referência variável: em uma frase, *ele* pode referir-se a Ricardo, em outra a Pedro, e assim por diante. Existem condições sintáticas que determinam a referência dos pronomes dentro da oração. Assim, por exemplo, em

(1) Ricardo disse que Pedro *o* beliscou.

podemos entender o pronome *o* como referindo-se a Ricardo, mas não a Pedro. Já em

(2) Ricardo disse que Pedro *se* penteia com manteiga.

o pronome *se* remete necessariamente a Pedro, nunca a Ricardo. Dizemos, então, que *o* e *Ricardo* em (1), assim como *se* e *Pedro* em (2), são **correferentes**.

As condições que governam a co-referência de pronomes com os demais SNs da oração são surpreendentemente complexas, e a tentativa de explicitá-las tem sido um dos grandes temas de pesquisa em sintaxe e semântica nos últimos anos. Nesta seção vou apresentar uma análise simplificada, que deve ser suficiente para dar conta da maioria dos casos do português.

11.1.1. Precedência; ciclo

É preciso, antes de mais nada, introduzir duas noções preliminares: a de precedência e a de aplicação cíclica.

Precedência significa simplesmente o fato de um elemento aparecer antes de outro na sentença; assim, em (1), *Ricardo* precede o pronome *o*, porque aparece antes dele na sentença.

Entende-se por **aplicação cíclica** a propriedade que têm certas regras de se aplicarem primeiramente às orações subordinadas, depois a suas principais. Assim, uma regra cíclica, ao se aplicar a uma estrutura, afeta primeiramente a oração mais subordinada de todas, aquela que não possui subordinadas; depois, aplica-se à oração imediatamente superior, aquela cuja única subordinada é a primeira mencionada; e assim por diante até chegar à oração principal da estrutura. Por exemplo, na frase

(3) Cristina sonhou que Tomás declarou que eu me candidataria a prefeito.

uma regra cíclica se aplicaria primeiro à oração *eu me candidataria a prefeito*; depois, à oração *Tomás declarou que eu me candidataria a prefeito*; finalmente, à oração *Cristina sonhou que Tomás declarou que eu me candidataria a prefeito.*

Como veremos logo a seguir, algumas regras de interpretação responsáveis pela determinação da referência dos pronomes dependem da relação de precedência e se aplicam ciclicamente às estruturas.

11.1.2. Regras de interpretação

As regras de interpretação que vamos estudar nesta seção são apenas três; mas, como se aplicam ciclicamente, a interpretação dos pronomes de uma estrutura pode envolver mais de três aplicações de regras. Isto é, as regras podem se aplicar uma vez no primeiro ciclo, mais uma vez no segundo etc. O resultado da aplicação das regras é a montagem gradativa de uma tabela de correferências, onde ficam tabulados os diferentes SNs presentes na estrutura, com suas relações de correferência.

A tabela faz parte da representação semântica da sentença. Por exemplo, à frase

(1) Ricardo disse que Pedro o beliscou.

corresponde a seguinte tabela de correferências:

Ricardo é correferente de *o* (ou: *Ricardo* <+cr> *o*);

Pedro não é correferente de *o* (ou: *Pedro* <–cr> *o*);

Ricardo não é correferente de *Pedro* (ou: *Ricardo* <–cr> *Pedro*).

A tabela, pois, exprime as relações de correferência existentes entre os diversos SNs de uma estrutura.

Passemos, então, às regras (começo a numeração com "R5", para manter continuidade com as quatro regras de interpretação já vistas no capítulo precedente):

Regra 5 (R5)

Marcar SN_1 *como correferente de* SN_2 *se* SN_2 *for um pronome reflexivo, e como não-correferente de* SN_2 *se* SN_2 *for um pronome não reflexivo, se as condições abaixo forem satisfeitas:*

SN_2 ainda não foi marcado, na tabela, como correferente de nenhum SN;
SN_1 pertence à oração principal do presente ciclo;
SN_1 precede SN_2;
SN_2 não é sujeito.

Esta regra é obrigatória e cíclica.

Regra 6 (R6)

Marcar SN_1 como correferente de SN_2 se SN_1 preceder SN_2 e SN_2 for um pronome.

Esta regra é opcional e cíclica.

Regra 7 (R7)

Marcar como não-correferentes *qualquer par de SNs para os quais não haja marca na tabela.*

Esta regra é obrigatória e se aplica depois da tabela pronta (isto é, não é cíclica).

Além dessas três regras, precisaremos de um filtro, a saber:

Filtro do reflexivo

Marcar como mal formada toda representação semântica que contiver um pronome reflexivo que não seja correferente de algum outro SN.

E vamos admitir, finalmente, que as marcas na tabela, uma vez feitas, não podem ser alteradas. Assim, se em algum ciclo as regras marcarem um pronome como correferente de algum SN, o mesmo pronome não poderá ser marcado como correferente de outro SN no ciclo seguinte; é fácil ver que isso introduziria inconsistência na tabela: o pronome *o* em (1) não poderia ser correferente de *Ricardo* e de *Pedro* ao mesmo tempo. Por outro lado, nada impede que um pronome já marcado como correferente de algum SN seja marcado, depois, como **não-correferente** de outro SN.

As regras R5 e R6 são, como vimos, cíclicas, e se aplicam na ordem dada, apenas uma vez em cada ciclo. Depois do último ciclo, aplica-se R7. Depois da aplicação das regras, o filtro marca como mal formadas as representações semânticas que incluem reflexivos não-correferentes de outro SN. Observe-se que isso só se aplica aos reflexivos, ou seja, um pronome não-reflexivo pode ficar sem atribuição de correferência na tabela.

Vejamos agora alguns exemplos da aplicação desse sistema à interpretação de frases. Comecemos com a frase

(4) Guilherme se penteou.

Aqui a questão do ciclo não se coloca, porque só há uma oração. Trata-se, portanto, apenas de estabelecer se *Guilherme* e o reflexivo *se* devem ser marcados como correferentes. *Guilherme* é o SN_1 mencionado na regra, e o reflexivo é o SN_2.

R5 marcará *Guilherme* e *se* como correferentes, obrigatoriamente, se as condições (a) a (d) forem satisfeitas. A condição (a) é satisfeita, porque inclusive ainda não se marcou nada na tabela; (b) é satisfeita

porque *Guilherme* pertence à oração principal do presente ciclo (aliás, a única oração); (c) também é satisfeita porque *Guilherme* precede *se*; e (d) é satisfeita porque *se* não é sujeito. Assim, marcamos na tabela

Guilherme <+cr> *se*

R7 não se aplica, pois todos os pares de SNs existentes (apenas um) constam da tabela.

Essa tabela concorda com nossa intuição, pois compreender (4) implica em entender *Guilherme* e *se* como se referindo à mesma pessoa.

Finalmente, o filtro inspecionará a tabela pronta e não achará violação, pois o único reflexivo existente recebeu marca de correferência.

Agora passemos ao exemplo

(5) Guilherme o penteou.

A única diferença entre (4) e (5) é que esta última tem um pronome não-reflexivo. Desse modo, todas as condições para a aplicação de R5 são satisfeitas; mas, como o pronome não é reflexivo, a regra marcará obrigatoriamente *Guilherme* e *o* como não-correferentes:

Guilherme <–cr> *o*

Novamente, o resultado é correto, pois concorda com o julgamento dos falantes: *Guilherme* e *o* em (5) não são a mesma pessoa.

Agora passemos a um caso mais complexo. Seja a frase

(6) Rui disse que Guilherme se acha um gênio.

Essa estrutura compreende duas orações: a subordinada *Guilherme se acha um gênio* e a principal *Rui disse que Guilherme se acha um gênio.* Assim, teremos que usar o ciclo: aplicaremos as regras primeiro à subordinada e depois à principal.

No primeiro ciclo (*Guilherme se acha um gênio*), aplicamos R5 à dupla *Guilherme/se.* As quatro condições são satisfeitas: *se* ainda não foi marcado como correferente de nenhum SN; *Guilherme* pertence à oração principal do presente ciclo; *Guilherme* precede *se*; *se* não é sujeito. E, como *se* é reflexivo, começamos a tabela com

Guilherme <+cr> *se*

Agora passamos ao segundo ciclo (*Rui disse que Guilherme se acha um gênio*); a dupla a ser examinada agora é *Rui/se.* R5 marcará os dois como correferentes se as condições forem satisfeitas. Mas isso não acontece, porque *se* já foi marcado como correferente de *Guilherme,* e portanto a condição (a) bloqueará nova atribuição de correferência. R5, portanto, não se aplica.

R6 tampouco se aplica, porque *se* já recebeu marca na tabela como correferente de *Guilherme,* e portanto não poderia ser igualmente correferente de *Rui.* R7 marca todos os pares de SNs não constantes da tabela como não-correferentes, a saber, *Rui/Guilherme* e *Rui/se,* acrescentando à tabela as marcas

Rui <–cr> *Guilherme*

Rui <–cr> *se*

O resultado, como se pode ver, concorda com o julgamento dos falantes.

Vejamos agora uma frase ligeiramente diferente de (6), a saber

(7) Rui disse que Guilherme o acha um gênio.

No primeiro ciclo, *Guilherme o acha um gênio,* R5 marca *Guilherme* e *o* como não-correferentes, porque *o* não é reflexivo e as condições (a)-(d) são satisfeitas. Assim, começamos a tabela com

Guilherme <–cr> *o*

R6 não se aplica, porque seu emprego implicaria em revisar a tabela, o que é interdito. Passamos então ao segundo ciclo, correspondente à oração principal: *Rui disse que Guilherme o acha um gênio.*

A dupla a ser interpretada agora é *Rui* e *o.* R5 não se pode aplicar porque a condição (a) não é satisfeita: *o* (= SN_2) já consta da tabela. Passamos a R6: como *o* é um pronome e *Rui* precede *o* (e, além disso, a dupla *Rui/o* ainda não foi interpretada), essa regra se aplica, produzindo

Rui <+cr> *o*

Desse modo, entende-se *o* como correferente de *Rui,* o que certamente é uma maneira de se interpretar (7).

Mas a regra R6 é definida como **opcional**, o que quer dizer que também é possível deixar de marcar *Rui* e *o* como correferentes. Nesse caso, R7 os marcará como não-correferentes, pois o par não figurará na tabela.

Em resumo, (7) corresponde a duas tabelas de correferência, a saber,

Guilherme <–cr> *o*

Rui <+cr> *o*

e

Guilherme <–cr> *o*

Rui <–cr> *o*

E, com efeito, (7) permite duas interpretações: na primeira, Rui afirma que Guilherme o acha, a ele Rui, um gênio. E na segunda a referência de *o* fica sem especificação: Rui afirma que Guilherme considera uma terceira pessoa um gênio. Isso significa que a opcionalidade de R6 deve ser mantida, a fim de permitir essa dualidade de interpretações.

Examinemos agora o exemplo

(8) Ele disse que Guilherme penteou Rui.

No primeiro ciclo, na verdade não há o que fazer, porque os dois SNs (*Guilherme* e *Rui*) não são pronomes e não poderiam ser correferentes. Passamos, então, diretamente ao segundo ciclo, onde há duas duplas de SNs a serem examinadas: *ele/Guilherme* e *ele/Rui* (desprezando-se, por razões óbvias, a dupla *Guilherme/Rui*). Mas observe-se que tanto R5 quanto R6 exigem, para sua aplicação, que o SN não-pronominal **preceda** o SN pronominal; e acontece que em ambas as duplas em exa-

me o pronome precede o SN não-pronominal. Isso quer dizer que nenhuma das duas regras poderá ser aplicada para estabelecer correferência entre *ele* e qualquer dos outros dois SNs.

O que vai acontecer é que não haverá marca nenhuma na tabela até o momento de aplicar R7, a qual, como sabemos, marcará todos os pares como não-correferentes:

ele <–cr> *Guilherme*

ele <–cr> *Rui*

Novamente, chegamos a um resultado que se adequa ao julgamento dos falantes: em (8), *ele* não se refere nem a Guilherme, nem a Rui, mas a uma terceira pessoa não-especificada.

Darei agora um exemplo para ilustrar a ação da condição (b) de R5:

(9) O gato que arranhou Rui também o mordeu.

O primeiro ciclo (*que arranhou Rui*) não tem nenhum SN pronominal tal como o definimos, e por isso nenhuma das regras se aplica. No segundo ciclo, há dois pares de SN a examinar: *o gato/o* e *Rui/o*. Vamos começar com o segundo par: o pronome *o* ainda não figura na tabela, vem depois de *Rui* e não é sujeito, portanto as condições (a), (c) e (d) são satisfeitas; mas *Rui* não pertence à oração principal do presente ciclo, e sim a uma subordinada, violando a condição (b). Portanto, R5 não pode aplicar-se a esses dois SNs; a dupla *Rui/o*, por enquanto, não dará entrada na tabela.

Continuando as tentativas de aplicação de R5, vamos examinar o par *o gato/o*. Aqui R5 pode aplicar-se: *o* ainda não figura na tabela, *o gato* pertence à oração principal do ciclo, *o gato* precede *o*, e *o* não é sujeito. Como *o* não é reflexivo, marcamos na tabela

o gato <–cr> *o*

Agora passamos à aplicação de R6. Como um dos dois pares já deu entrada na tabela, só precisamos examinar *Rui/o* — dupla que, conforme nos lembramos, ainda não consta da tabela. Como *Rui* precede o pronome, podemos (opcionalmente) marcar

Rui <+cr> *o*

E, com efeito, (9) pode ser entendido como dizendo que Rui, além de arranhado, também acabou mordido. Alternativamente (já que R6 é opcional) pode-se entender que as mordidas foram reservadas a uma outra pessoa, cujo nome não se menciona. (Na maioria dos casos, evidentemente, essa ambigüidade do pronome é desfeita pela intervenção de elementos extralingüísticos do contexto; por exemplo, posso enunciar (9) apontando para Bruno, de maneira que se entenderá *o* como referindo-se a Bruno.)

Quando uma questão é tão complexa como esta, exemplificação

nunca é demais. Assim, vamos examinar mais um caso:

(10) Carvalho informou a Almeida que ele estava despedido.

No primeiro ciclo (*ele estava despedido*), não há o que fazer, porque só há um SN, *ele*. Passamos ao segundo ciclo, que compreende toda a frase. R5 não se aplica, porque o pronome *ele* é sujeito, violando assim a condição (d). R6 pode aplicar-se à dupla *Carvalho/ele* ou então à dupla *Almeida/ele* (já que em ambos os casos o pronome vem depois do SN não-pronominal); mas R6 não pode aplicar-se a **ambas** as duplas, porque cada regra só se aplica uma vez por ciclo. O resultado é que teremos (após a aplicação de R7) pelo menos duas interpretações, correspondentes às tabelas

Carvalho <+cr> *ele*
Almeida <–cr> *ele*

ou então

Carvalho <–cr> *ele*
Almeida <+cr> *ele*

Em outras palavras, em (10) o despedido pode ser Carvalho ou Almeida; creio que isso corresponde adequadamente à intuição dos falantes.

Além disso, há uma outra interpretação possível, visto que R6 é de aplicação opcional: se não a aplicarmos, *ele* não será marcado como correferente de nenhum outro SN, e R7 marcará os três SNs *Carvalho, Almeida* e *ele* como tendo referências independentes. Isso corresponde a entender (10) como afirmando que Carvalho informou a Almeida que uma terceira pessoa estava despedida, o que, certamente, é uma das possibilidades de interpretação de (10).

11.1.3. Condições de consistência

Complexo como é, o sistema acima exposto ainda assim não é suficiente para explicar todos os fatos. As regras R5, R6 e R7, que constituem o essencial do mecanismo de interpretação dos pronomes, são basicamente formais; ou seja, aplicam-se à estrutura sintática da sentença, derivando daí a tabela de correferência.

No entanto, há também condições propriamente semânticas que precisam ser levadas em consideração ao se estabelecerem as relações de correferência. Por exemplo, se submetermos às regras vistas a frase

(11) Guilherme disse que eu a odeio.

será possível derivar uma tabela onde *Guilherme* e *a* são marcados como correferentes (ao lado de outra tabela na qual não o são, em virtude da opcionalidade de R6). Mas a intuição dos falantes não aceita essa interpretação, pela razão muito simples que *Guilherme*, nome de homem, não po-

deria referir-se à mesma pessoa que o feminino *a*.

Essa razão, evidentemente, não tem nada a ver com a estrutura sintática da sentença; por isso, vamos chamá-la de **condição semântica de consistência**, ou mais simplesmente **condição de consistência**. No caso, a formulação da condição é fácil; pode ser algo como

marcar como mal formada uma representação semântica se houver divergência de gênero, pessoa e/ou número entre dois SNs correferentes.

(Essa condição funciona dentro dos limites do período; no discurso em geral, é possível colocar em correferência SNs divergentes, como, por exemplo, *Guilherme* e *a vítima.*)

As condições de consistência inspecionam cada passo da montagem da tabela, não tendo, portanto, um lugar especificado na ordenação. Não são, na verdade, regras nem filtros, mas condições de boa formação das tabelas de correferência.

Voltando ao exemplo (11), sabemos que as regras permitem duas tabelas alternativas, ou seja,

Guilherme <+cr> *a*

ou

Guilherme <–cr> *a*

A primeira dessas tabelas deverá ser marcada como mal formada e, portanto, nunca ocorrerá. Os falantes só reconhecem a segunda, na qual o pronome se refere a alguma outra pessoa que não Guilherme.

É de se esperar que existam muitas dessas condições de consistência, porque as frases não apenas devem ser bem estruturadas, mas ainda devem ter um significado coerente. O estudo de tais condições é parte da semântica, mas nesta *Gramática* não será possível abordá-lo; deveremos contentar-nos com alguns exemplos.

A condição de consistência que acabamos de ver não é a única condição semântica relevante para o estabelecimento da referência dos pronomes. Vou dar agora mais um exemplo de como restrições de significado podem afetar a compreensão das relações de correferência.

Vimos na seção precedente que a representação semântica da frase

(10) Carvalho informou a Almeida que ele estava despedido.

inclui três possibilidades de interpretar a referência de *ele*: pode ser Carvalho, pode ser Almeida, ou pode ser uma terceira pessoa. Essas três possibilidades são todas previstas pela análise.

Agora tomemos a frase

(12) Carvalho ordenou a Almeida que ele lavasse o carro.

Essa frase não é totalmente aceitável para todos os falantes; mas uma boa parte deles a aceita, na acepção em que *ele* é correferente de *Almeida* (creio que mesmo quem

considera a frase algo marginal pode perceber isso com clareza). Ora, essa interpretação vai contra a análise, tal como foi desenvolvida até agora, porque (12) é estruturalmente semelhante a (10), e portanto *ele* deveria poder referir-se a Carvalho, a Almeida ou ainda a uma terceira pessoa. Como explicar essa diferença semântica entre duas frases tão parecidas?

Aqui temos de lançar mão novamente de uma condição semântica. Parece intuitivamente correto afirmar que o verbo *ordenar* admite um destinatário da ordem (ou seja, um constituinte com o papel semântico de **meta**), mais a expressão da coisa que deve ser feita. Além disso, o verbo exige que a meta e o agente da coisa a ser executada sejam correferentes. Em outras palavras, não faria muito sentido ordenar a Almeida algo que deve ser executado por Rodrigues. Isso é parte da matriz semântica de *ordenar,* ou seja, parte de seu significado. Há outros verbos que fazem a mesma exigência quanto à relação entre meta e a coisa a ser executada: *mandar, dizer* (no sentido de "ordenar") e outros.

Podemos, então, utilizar esse traço semântico do verbo *ordenar* para explicar por que nem todas as interpretações atribuídas pelas regras R5-R7 à frase (12) se realizam. O traço semântico mencionado pode ser expresso da seguinte maneira:

Traço semântico *(presente na matriz de verbos como ordenar, mandar e outros): Exige correferência entre a* meta *da oração principal e o* agente *da subordinada.*

Traços como esse se denominam **redes de correferência**. Uma convenção geral estabelece que qualquer frase que contenha determinado verbo deve receber uma interpretação semântica que satisfaça a rede de correferência associada àquele verbo. Caso contrário, marca-se a frase como mal formada.

Agora podemos dizer por que (12) só tem uma das três interpretações esperadas: como *Almeida* representa a meta da oração principal, *ele* deve ser correferente de *Almeida.* Tanto *Carvalho* quanto uma terceira pessoa não-especificada ficam excluídos, por não estarem desempenhando o papel semântico de meta na oração principal.

Observe-se que em (10) nada impede que o pronome seja marcado como correferente de *Carvalho* ou então de *Almeida,* porque o verbo *informar* não tem a rede de correferências dada acima. E, com efeito, é bastante claro que quando A informa algo a B, essa informação pode referir-se a qualquer pessoa: A, B ou outra pessoa qualquer.

Nas seções acima, dei uma idéia de como o componente semântico pode funcionar para atribuir a cada pronome uma referência.

11.2. ELIPSES

11.2.1. O problema

Vamos passar agora a um tópico relacionado com o da interpretação da referência dos pronomes, mas não idêntico a ele: a interpretação de elementos elípticos.

O problema pode ser ilustrado por uma frase como

(13) Pedrão declarou que renunciaria.

A interpretação dessa frase contém a ação de "declarar", cujo agente é Pedrão; esses ingredientes semânticos se derivam facilmente a partir de elementos explícitos na própria sentença, com o auxílio das regras R1-R3 (ver a seção 10.2.3.). Mas, além disso, a interpretação de (13) também contém a ação de "renunciar", atribuída ao mesmo agente Pedrão. E, como *Pedrão* não é sujeito explícito de *renunciaria* (na verdade, pela análise sintática desenvolvida no capítulo 3, esse verbo não tem sujeito), surge o problema de como atribuir a *renunciaria* um agente idêntico ao de *declarou.*

Tradicionalmente, fala-se de um "sujeito elíptico", "oculto" ou "subentendido" em orações como a subordinada de (13). Segundo nossa análise, evidentemente, não há sujeito nenhum ali; mas é inegável que há algo "subentendido". Apenas, sustentaremos que se trata de uma entidade semântica e não sintática: o mecanismo de interpretação aventado nas seções seguintes parte desse pressuposto. Assim, a subordinada *renunciaria* será analisada (sintaticamente) como oração sem sujeito, nos termos da seção 3.2.2.3.; mas a interpretação semântica dessa subordinada deverá conter um **agente**, que é o mesmo Pedrão que funciona como agente da principal.

O problema, pois, se resume à necessidade de elaborar um sistema de regras que permita interpretar a oração *renunciaria,* que não tem sujeito, como tendo um agente.

A questão da interpretação de elementos elípticos não se limita a sujeitos, nem a SNs. Assim em

(14) Maria picou o pato e pôs na panela.

temos um "objeto direto elíptico" [essa é evidentemente uma designação cômoda, mas inexata, pois não há objeto direto na segunda oração de (14)]. E em

(15) Serafim toca flautim, e Ivone trombone.

temos na segunda oração um "verbo (NdP) elíptico", cujo significado é "toca". Esses casos também serão discutidos nas seções seguintes.

11.2.2. Semântica do sufixo de pessoa-número

Vamos ocupar-nos, primeiramente, do caso dos "sujeitos ocultos" tradicionais, como em

(13) Pedrão declarou que renunciaria.

Nosso ponto de partida será o seguinte: em uma frase como

(16) Eu comprei uma espora nova.

a informação sobre o agente da ação de comprar (isto é, "eu") é redundante: provém de duas fontes, a saber, (a) o sujeito *eu* e (b) a forma do verbo, em particular o sufixo de pessoa-número. Por ser redundante, a frase poderia, em princípio, prescindir de um dos dois elementos, sem perder informação; e com efeito a língua permite, neste caso, a omissão do sujeito:

(17) Comprei uma espora nova.

Essa frase é, para todos os efeitos, sinônima de (16).

Ora, isso significa que as regras de interpretação precisam ter condições de extrair de (17) a informação de que o agente de *comprei* é "eu". Note-se que em (16) isso não representa problema, já que *eu* é o sujeito e as regras R1-R3 (seção 10.2.3.) atribuirão a esse termo o papel semântico de agente. Mas (17) não tem sujeito, de modo que as regras propostas até o momento não saberiam processá-la de modo a atribuir corretamente o papel de agente. Será necessário formular alguma regra ou princípio novo, que possa lidar com casos como o de (17).

Teremos, na verdade, de fazer com que o sufixo de pessoa-número forneça a mesma informação que o sujeito. Isso pode ser efetuado pela regra seguinte:

Regra 8 (R8)

O sufixo de pessoa-número vale, para efeito da aplicação de regras semânticas, como um sujeito pronominal, *marcado com os traços de pessoa e número do sufixo.*

(Nesses casos, por comodidade, falarei de **sujeito vazio**.)

Note-se, primeiro, que essa equivalência entre um sujeito pronominal e a ausência de sujeito, em casos de verbo marcado para pessoa, só existe no plano semântico. Sintaticamente, orações com sujeito (pronominal ou não) se analisam de maneira diferente de orações sem sujeito.

Conforme veremos, a postulação de R8 permite interpretar corretamente as orações sem sujeito sem necessidade de acrescentar mais nada (isto é, utilizando-se as regras já formuladas até o momento).

Vamos então ver como ficará a interpretação de (17) utilizando-se R8, mais as outras regras relevantes:

(17) Comprei uma espora nova.

Vamos aplicar as regras 1 a 3, dadas na seção 10.2.3. R1 atribui ao objeto direto o papel de paciente; assim, *uma espora* se interpreta como "a coisa comprada". R2 não se aplica, pois não há atributo; R3 interpreta o sujeito como agente, instrumento ou paciente, nessa ordem de preferências, e de acordo com os traços semânticos do verbo e do próprio sujeito. Ora, *comprar* admite agente, e portanto o sujeito seria interpretado como agente. Não há sujeito, mas há sufixo de pessoa-número (primeira pessoa do singular), que de acordo com R8 equivale, para efeitos de interpretação, a um sujeito pronominal de primeira pessoa do singular. Dessa maneira, introduz-se na representação semântica de (17) um agente "eu". O resultado é uma interpretação semântica correta: agente "eu", paciente "uma espora nova".

Já na frase

(16) Eu comprei uma espora nova.

o mesmo processo pode ser aplicado, dando o mesmo resultado. Mas, além disso, há uma outra fonte (redundante) da qual se pode derivar a identidade do agente, a saber, o sujeito *eu*. O resultado é, novamente, o mesmo de (17). Pode-se dizer que (16) tem duas interpretações semânticas (é, de certo modo, ambígua); mas essas duas interpretações são idênticas — o que, na prática, elimina a ambigüidade.

Pode-se perguntar por que não se evita a menção ao sujeito, formulando R8 de maneira a derivar diretamente o agente: o sufixo de pessoa-número teria a interpretação de agente.

Não se pode fazer isso porque o sufixo fornece sempre o papel semântico que o sujeito teria, se estivesse presente, mesmo quando o sujeito tem algum outro papel que não o de agente. Assim, em

(18) Eu apanhei de minha mulher.

o sujeito é paciente. E "eu" é paciente também em

(19) Apanhei de minha mulher.

Desse modo, fica claro que o que o sufixo denuncia não é o agente, mas o sujeito; isso justifica a maneira como formulei R8.

Uma última observação: a análise acima fornece um argumento, de natureza semântica, em favor da análise da frase

(17) Comprei uma espora nova.

como oração sem sujeito. O argumento é o seguinte: se (17) tiver sujeito (oculto, no caso), não será possível distinguir seu caráter não-redundante, em face de

(16) Eu comprei uma espora nova.

que é redundante no que se refere à expressão do agente. Se (17) também tiver sujeito, as duas frases serão essencialmente idênticas para efeito de interpretação semântica;

conseqüentemente, na interpretação de (17) aplicaremos não apenas a Regra 8, mas também a Regra 3 (vista em 6.2.2.3.), que como vimos serve para interpretar o sujeito. Em outras palavras, a informação sobre o agente (que é o papel semântico do sujeito) será redundante em (17), pois provém de duas fontes: do sufixo de pessoa-número e do sujeito. Mas isso vai contra a intuição de que (16) é claramente redundante em um sentido em que (17) não é. Em outras palavras, é bastante claro que em (17) a informação sobre o agente provém de uma só fonte, e em (16) de duas. Isso poderá ser representado na semântica se adotarmos a análise segundo a qual (17) não tem sujeito.

A alternativa seria tratar o sujeito oculto diferentemente do sujeito claro; algo como dizermos que o sujeito oculto escapa à interpretação semântica. Mas essa solução prejudica a motivação básica para a postulação de sujeitos ocultos, que é justamente a possibilidade de tratá-los da mesma maneira que os sujeitos claros. O sujeito oculto já difere do sujeito claro formalmente (um é explicitado, o outro não); se formos distinguir os dois também semanticamente, que razão sobra para incluir sujeitos ocultos na análise? Em suma, acredito que temos aqui mais um argumento em favor de analisar (17) e frases semelhantes como sem sujeito.

11.2.3. Semântica do sufixo de 3.ª pessoa

Antes de encerrar a discussão dos efeitos da regra de interpretação do sufixo de pessoa-número, vou abordar um assunto importante relativo a uma diferença entre a terceira pessoa e as demais pessoas gramaticais. Trata-se, em resumo, da vaguidão de referência que caracteriza a terceira pessoa e que acaba resultando em inaceitabilidades um tanto inesperadas.

Vamos examinar primeiro a terceira pessoa do singular. Observa-se que frases com verbo na terceira pessoa do singular e sem sujeito são em geral de aceitabilidade muito baixa. Assim, embora se possa perfeitamente dizer

(17) Comprei uma espora nova.

a frase

(20) ?? Comprou uma espora nova.

pelo menos isoladamente, é muito menos aceitável. A frase (20) só se aceita se ocorrer inserida em um contexto que permita especificar o agente, como em

(21) Meu pai foi à cidade e comprou uma espora nova.

A razão, parece-me, é que a informação fornecida pelo sufixo de pessoa-número é menos especificada do que a fornecida por um sujeito

explícito — especialmente no caso da terceira pessoa do singular. Mesmo um pronome como *ele* fornece informação mais rica do que o sufixo sozinho: *ele* é masculino e se refere a uma pessoa diferente do ouvinte. Um simples sufixo de terceira pessoa do singular poderia referir-se a "ele", "ela" e ainda a "você", por exemplo. Ao que parece, uma forma verbal de terceira pessoa do singular é tão pouco informativa a respeito do sujeito (vazio) que, em condições de isolamento, chega a tornar-se inaceitável. Esse tipo de problema foi definido por Grice (1975) como uma violação do **princípio de cooperação**, espécie de contrato tacitamente assumido pelos interlocutores de um diálogo. No caso, a parte do princípio que é violada é a máxima de quantidade, assim formulada por Grice:

> Faça sua contribuição tão informativa quanto necessário para os atuais objetivos do diálogo. [apud Levinson, 1983, p. 101]

Uma frase como (20) é deficiente desse ponto de vista: falta informação suficiente quanto ao agente da ação de "comprar". Sua inaceitabilidade, portanto, é de natureza pragmática, antes que semântica ou sintática. Em um trabalho anterior, levantei a hipótese de que este seria um fenômeno sintático (ver Perini, 1989, p. 82); mas agora acho que a solução dada aqui é bem mais plausível.

Vejamos como isso se aplica a um caso mais complexo:

(22) Patrícia acha que ganha pouco.

Uma vez que aceitamos R8 como parte da análise, será possível aplicar as regras que atribuem referência aos pronomes também a essa frase, já que o sufixo de pessoa-número equivale a um sujeito pronominal. Assim, é como se a oração subordinada de (22) contivesse um sujeito pronominal; esse sujeito pronominal (vazio) pode então ser submetido à aplicação de R5 a R7. O resultado deveria ser, evidentemente, idêntico ao que se obteria se houvesse um pronome explícito antes do verbo, isto é,

(23) Patrícia acha que ela ganha pouco.

Assim, vamos aplicar a (22) as regras R5-R7. R5 não se aplica, porque o pronome vazio é sujeito; R6, então, poderá marcar *Patrícia* e o pronome vazio como correferentes, produzindo uma interpretação correta.

Mas como R6 é opcional, pode deixar de se aplicar; nesse caso, R7 se aplicará, marcando *Patrícia* e o sujeito vazio como não-correferentes. Isso é o que acontece com (23) e deveria também acontecer com (22). Mas aqui surge uma dificuldade: (22) não tem essa outra interpretação; ou seja, só se pode entender que Patrícia acha que ela mesma, Patrícia, ganha pouco, e não que Patrí-

cia acha que outra pessoa ganha pouco. Não se verifica a identidade esperada de interpretação entre (22) e (23).

A explicação é paralela à que foi dada para a inaceitabilidade de

(20) ?? Comprou uma espora nova.

A frase (22) teria, em princípio, duas interpretações, tal como (23). Mas uma dessas interpretações, aquela em que R6 não se aplicou, inclui um sufixo de terceira pessoa do singular, sem sujeito (ou, se preferirem, com sujeito vazio); e esse sujeito vazio não é marcado como correferente de nenhum SN. O que resulta é uma indefinição excessiva da referência do sujeito de *ganha*. A aceitabilidade da frase, nessa acepção, é muito baixa, e o falante tem a impressão nítida de que a frase não é ambígua — um efeito da preponderância muito grande da outra acepção, em que o sujeito vazio é correferente de *Patrícia*. O problema, obviamente, não surge com (23), pois ali o sujeito da subordinada (*ela*) é explícito e bem mais informativo do que o sujeito vazio de (22).

Passemos agora à terceira pessoa do plural. Há uma diferença importante entre o singular e o plural, porque a terceira pessoa do plural sem sujeito pode ser interpretada como tendo "agente não-especificado", por efeito da Regra 4 (seção 10.2.6.), a saber,

Regra 4 (R4)

Introduzir "agente (não-especificado)" na representação semântica de frases com NdP na 3ª pessoa do plural, sem sujeito.

Esta regra é opcional.

R4, como o leitor deverá estar lembrado, é responsável pela interpretação de frases como

(24) Quebraram meu relógio.

com agente não-especificado.

Quando da formulação da regra, avisei que as razões de sua opcionalidade seriam explicadas oportunamente. Agora já temos condições de fazê-lo.

Em uma frase como (24), a opcionalidade de R4 não parece relevante, porque (24) não é ambígua. No entanto, não é ambígua porque uma das acepções possíveis, aquela que não envolve a aplicação de R4, é mal formada pelas razões acima vistas em conexão com (20): insuficiência de informação acerca da referência do sujeito.

Mas se tomarmos a frase

(25) Meus irmãos dizem que bateram na porta.

a situação muda. Essa frase é claramente ambígua: (a) meus irmãos dizem que alguém (não-especificado) bateu na porta, ou então (b) meus irmãos dizem que eles mesmos bateram na porta. Isso pode ser entendido como decorrente da análise, da

seguinte maneira: em primeiro lugar, submetemos a oração subordinada de (25) à aplicação de R4. Caso essa regra se aplique, *bateram* terá um agente não-especificado, e a acepção (a) será produzida. Se R4 (que é opcional) não se aplicar, então R8 atribuirá a *bateram* um sujeito vazio, que acabará sendo marcado, no ciclo da oração principal, como correferente de *meus irmãos*, gerando a acepção (b).

Evidentemente, é necessário impedir que R8 se aplique quando R4 já se aplicou, pois, se isso acontecer, a interpretação (a) nunca surgirá. Mas essa possibilidade é bloqueada pelo princípio geral de que marcas feitas na tabela não podem ser modificadas. A identificação do agente de *bateram* como não-especificado constaria da tabela, da seguinte maneira:

[suj. vazio de *bateram*] <+cr> "não-especificado".

Vimos em todos esses exemplos como o acréscimo de R8 permite a interpretação de frases contendo sujeitos elípticos, utilizando-se as regras gerais de construção da tabela de correferências.

Isso vale em geral para as frases sem sujeito (e com verbo flexionado), e não é necessário multiplicar os exemplos. O leitor poderá por si mesmo verificar o funcionamento do sistema analisando outras frases. A análise aqui exposta não é perfeita, embora seja suficiente para descrever a maioria dos casos.

11.2.4. Semântica do infinitivo impessoal

11.2.4.1. Regra de interpretação do infinitivo

Agora vamos considerar casos de orações sem sujeito e sem verbo flexionado, como

(24) Maria espera vencer a partida.

A oração *vencer a partida* não tem sujeito e o verbo não tem sufixo de pessoa-número. Essas frases apresentam um novo problema de interpretação semântica; por exemplo, em (26), *Maria* se entende como agente de *vencer*, embora seja, sintaticamente falando, apenas o sujeito de *espera*. Nesta seção, procurarei desenvolver uma análise que descreva a interpretação semântica de frases com infinitivo impessoal, como (26).

Examinando apenas (26), a regra parece ser simples: quando um verbo não tem sujeito nem sufixo de pessoa-número, vale como sujeito o primeiro SN à sua esquerda. No caso, *Maria* é o primeiro SN à esquerda de *vencer*, e portanto funcionaria como seu sujeito (chama-se a esse SN o **controlador** do papel semântico que seria atribuído ao sujeito de *vencer*; assim, diremos que *Maria* **controla** o sujeito vazio de *vencer*).

Mas as coisas não são assim tão simples. Em certos casos, parece que

o controlador é o **segundo** SN à esquerda, como em

(27) Maria prometeu a Bruno fazer o café.

Entende-se que o agente de *fazer* é Maria, e não Bruno.

Em outros casos, o SN controlador se encontra **à direita** do infinitivo, como em

(28) Perder dinheiro não incomoda Bruno.

Essa frase se interpreta como contendo o ingrediente "Bruno perder dinheiro", embora *Bruno* esteja à direita de *perder*.

Na verdade, o fator relevante aqui não é a posição linear relativa do infinitivo e do SN controlador. Em todos os exemplos vistos, há uma coisa em comum: o SN controlador está sempre em uma oração que contém aquela onde está o infinitivo; ou seja, o infinitivo está sempre em uma oração subordinada àquela onde está o SN controlador. Isso terá de ser levado em conta quando investigarmos as condições de identificação do SN controlador de um infinitivo.

Mas primeiro vou formular a regra de interpretação propriamente dita, a saber,

Regra 9 (R9)

Em casos de verbo no infinitivo impessoal (sem sujeito), o papel semântico atribuído pelo verbo ao sujeito é preenchido com uma cópia da matriz semântica do SN controlador.

Regra obrigatória.

Vejamos como R9 funciona no caso de

(26) Maria espera vencer a partida.

O verbo *vencer* está aí sem sujeito, no infinitivo impessoal. Esse verbo, em virtude de seus traços semânticos (isto é, de seu significado) atribui ao sujeito o papel semântico de **agente**. A aplicação de R9 significa que a interpretação semântica da frase deve conter (a matriz semântica de) *Maria* como agente de *vencer*. Note-se que *vencer* continua sem sujeito, já que R9 só afeta o nível da representação semântica, e não a estrutura sintática da sentença. Pode-se entender o efeito de R9 como estipulando que a matriz semântica de *Maria* vale tanto para *espera* quanto para *vencer*.

O mesmo vai acontecer, claro, com *Maria* e *fazer* em (27), e com *Bruno* e *perder* em (28). A aplicação da regra, em si, é pois muito simples.

O problema, evidentemente, está em determinar qual é o SN controlador, dentre os vários candidatos que a oração principal pode conter. Esse é um dos problemas clássicos da semântica moderna, e não se pode dizer que esteja cabalmente resolvido. A seguir darei uma solução parcial, suficiente para descrever a grande maioria dos casos; mas fique claro que há casos problemáticos que não entram nessa solução.

11.2.4.2. O problema do controle

A questão de determinar o SN controlador se denomina "o problema do controle". Já vimos parte da solução acima, quando mencionei o fato de que o SN controlador se encontra sempre em uma oração principal em relação àquela onde está o infinitivo. Agora vamos examinar a questão em maior detalhe.

O SN controlador deve estar não apenas em uma oração principal, mas também na **primeira oração**, logo acima daquela onde está o infinitivo. Isto é, se um SN estiver em uma oração (O_1) que contém uma segunda (O_2), que por sua vez contém uma terceira (O_3), que é onde existe um infinitivo, esse SN não poderá controlar o infinitivo. Essa situação ocorre na frase

(29) Maria disse a Bruno que perder a corrida levaria Ayrton ao desespero.

Vamos primeiro dividir (29) em orações. A principal é

[*Maria disse a Bruno que perder a corrida levaria Ayrton ao desespero*]

Dentro dessa principal, há uma subordinada que é

[*perder a corrida levaria Ayrton ao desespero*]

E essa subordinada, por sua vez, tem uma subordinada que é

[*perder a corrida*]

Resumindo, o período se analisa como

[Mª disse a B que {(perder a corrida) levaria A ao desespero}]

Como se vê, o infinitivo está em O_3, e há nas outras orações três SNs que poderiam, em princípio, ser seus controladores: *Ayrton, Maria* e *Bruno.* Mas é fácil ver que somente *Ayrton* funciona como controlador: a frase (29) não menciona a possibilidade de Maria nem Bruno perderem a corrida; o único candidato é Ayrton, tanto é que a frase não é ambígua quanto a quem poderia perder a corrida.

Desse modo, podemos começar a delimitar o "SN controlador" mencionado em R9 assim:

Um SN só pode controlar um infinitivo que esteja em uma oração imediatamente subordinada àquela em que se encontra o SN.

Isso quer dizer que se a oração intermediária (O_2 em (29)) não contiver nenhum SN, ainda assim um SN em O_1 não poderá controlar o infinitivo. Isso se vê no exemplo

(30) Maria disse a Bruno que comer caju faz mal à vista.

Comer caju é imediatamente subordinada a *comer caju faz mal à vista.* Mas, nesse ciclo, não há nenhum SN que possa ser controlador de comer: *caju*, além de semanticamente inadequado, está na mesma oração que o infinitivo, e portanto não pode con-

trolá-lo; e *vista* é igualmente excluído por razões semânticas óbvias.

No ciclo seguinte, o da oração *Maria disse a Bruno que comer caju faz mal à vista*, há dois SNs elegíveis, *Maria* e *Bruno*. No entanto, nenhum deles controla *comer*; por isso, entende-se o agente de *comer* como indeterminado, algo como "qualquer pessoa comer caju" (parece um fato geral que um infinitivo não-controlado se interpreta como tendo agente — ou o que seja — indeterminado).

Assim podemos complementar R9 com a definição do que vem a ser um "SN controlador"; por ora, a definição pode ser a seguinte:

Definição (parcial) de "SN controlador": um SN é controlador de um infinitivo quando está em uma oração à qual a oração que contém o infinitivo é imediatamente subordinada.

Essa definição é parcial porque falta especificar o efeito de traços idiossincráticos dos diferentes verbos. Passo agora a um breve estudo de alguns desses traços.

11.2.4.3. Traços semânticos e controle

Quando estudamos a interpretação dos pronomes na seção 11.1., vimos que a mesma é condicionada em parte por traços semânticos dos verbos. Assim, uma frase cujo verbo principal é *ordenar* (ou *mandar*) exige que haja correferência entre a meta da oração principal e o agente da subordinada. Isso é parte do significado de *ordenar*, e explica por que não podemos entender *Carvalho* como correferente de *ele* em

(12) Carvalho ordenou a Almeida que ele lavasse o carro.

Já sabemos que traços desse tipo recebem o nome de **redes de correferência**; as redes de correferência continuam válidas, evidentemente, em frases com infinitivos. Vamos examinar a frase

(31) Maria mandou Bruno fazer o café.

As duas orações são [*Maria mandou Bruno fazer o café*] e [*fazer o café*]. R9 nos permitiria, em princípio, escolher *Maria* ou *Bruno* como controlador de *fazer*. No entanto, se o SN *Maria* for escolhido, a interpretação semântica resultante será mal formada, porque a meta de *mandou* (Bruno) não é correferente do **agente** de *fazer* (Maria). A outra interpretação possível, com *Bruno* como controlador, é bem formada, já que *Bruno* é meta da oração principal. Isso explica por que, apesar do que estipula R9, a frase (31) não é ambígua, e apenas *Bruno* pode ser entendido como agente de *fazer*.

Um verbo que tem comportamento diferente do de *mandar* é *prometer*. Esse verbo exige que o agente (que é concomitantemente fonte da promessa) da principal seja correfe-

rente do agente da subordinada. Voltemos ao exemplo

(27) Maria prometeu a Bruno fazer o café.

Essa frase é sintaticamente muito semelhante a (31), e a aplicação de R9 deveria admitir que tanto *Maria* quanto *Bruno* controlassem *fazer*. Mas o traço semântico mencionado de *prometer* evita que *Bruno*, que não é agente (fonte) de *prometer*, seja o controlador de *fazer*; apenas *Maria* (agente e fonte) pode sê-lo. E, com efeito, a representação semântica de (27) inclui "Maria" como agente de *fazer*.

O verbo *pedir* apresenta um terceiro tipo de exigência. Em uma frase como

(32) Maria pediu a Bruno para sair.

podemos entender tanto *Maria* quanto *Bruno* como controladores, e portanto como agentes de *sair*. Exprime-se isso atribuindo ao verbo *pedir* a exigência de que o agente da subordinada seja correferente do agente (fonte) ou então da meta da principal.

11.2.4.4. Infinitivos com agente indeterminado

A maneira como formulamos R9 impede a regra de aplicar-se quando a oração imediatamente acima daquela que contém o infinitivo não tem nenhum SN compatível com o papel semântico que o verbo no infinitivo atribui ao sujeito. É o caso da frase (30), que já vimos acima:

(30) Maria disse a Bruno que comer caju faz mal à vista.

Conforme foi apontado na seção 11.2.4.2., *comer* fica com agente indeterminado. O mesmo acontece em uma versão mais simples de (30), a saber,

(33) Comer caju faz mal à vista.

Aqui, novamente, R9 não pode aplicar-se, por falta de um SN compatível na oração principal; o único SN aí presente, *vista*, não poderia ser interpretado como agente de *comer* por razões de consistência semântica.

Mas se substituirmos *vista* por um nome de pessoa, as coisas mudam:

(34) Comer caju faz mal a Antônio.

Agora nenhuma condição de consistência impedirá R9 de selecionar *Antônio* como controlador de *comer*. E, realmente, a representação semântica de (34) inclui "Antônio" como agente de *comer*. Como se vê, a diferença semântica existente entre as subordinadas de (33) e (34) (que são sintaticamente semelhantes) se deve à ação da restrição que impede que se atribua a um SN de significado inanimado o papel semântico de agente.

Naturalmente, deveremos acrescentar alguma coisa à Regra 4, que vimos em 10.2.6. O leitor talvez se lembre que essa regra é responsável pela introdução de "agente não-especifi-

cado" na representação semântica de frases com NdP na terceira pessoa do plural, sem sujeito. Agora teremos de dizer que também um infinitivo impessoal não-controlado é interpretado como tendo agente não-especificado (há alguns problemas ainda a resolver quanto a R4, mas prefiro não entrar em uma discussão que poderia levar-nos muito longe). De qualquer forma, como já notei brevemente acima, as condições em que o português permite a introdução de um agente não-especificado nas representações semânticas ainda estão por ser pesquisadas em detalhe.

11.2.5. Elipse em estruturas coordenadas

Todos os exemplos de recuperação de elementos elípticos vistos acima ocorrem dentro de sentenças contendo subordinação. Em estruturas coordenadas, o sistema visto não funciona, mesmo porque não há possibilidade de se aplicar a definição de "SN controlador" dada na seção 11.2.4.2. Nesta seção, tentaremos estabelecer as condições que governam a recuperação de elementos subentendidos em estruturas coordenadas. Não farei nenhuma tentativa de generalizar a análise, procurando unificar as duas situações (subordinação e coordenação), embora isso seja, evidentemente, um tema importante de investigação. Aqui, como em geral nesta *Gramática*, nosso objetivo principal é descrever os fatos da língua, limitando a um mínimo o esforço de generalização.

11.2.5.1. Orações coordenadas sem sujeito

O caso mais simples a ser examinado é o de orações coordenadas sem sujeito; essas orações têm sempre o verbo em forma finita, o que vai facilitar a análise. Um exemplo pode ser

(35) César impressionava os homens e beliscava as mulheres.

O problema aqui é a identificação do agente de *beliscava*, isto é, da matriz semântica a ser colocada no lugar selecionado pelo verbo para o sujeito. Não havendo sujeito, é preciso preencher esse lugar através de regras de interpretação.

A regra 8, vista na seção 11.2.2., se aplica à segunda oração de (35), de modo que o sufixo equivale a um sujeito pronominal de terceira pessoa do singular. Nosso problema se reduz, portanto, a encontrar o antecedente desse pronome, ou seja, o SN que o controla. A definição já vista não nos serve, pois só se aplica a sentenças compostas por subordinação. Como veremos, aliás, parece que as estruturas coordenadas requerem a formulação de um sistema especial de regras, diferentes em vários aspectos das que interpretam as estruturas compostas por subordinação.

No caso do sujeito, a regra é a seguinte:

Regra 10 (R10)

Em orações coordenadas sem sujeito, o NdP é controlado pelo sujeito da primeira oração à esquerda.

Essa regra permite a interpretação de *beliscava* em (35) como exprimindo uma ação praticada por César. Se houver outro SN na primeira oração, este não poderá controlar *beliscava*, porque não é o sujeito. Assim, em

(36) César chamou Pompeu e dispensou Crasso.

somente *César* é que pode ser entendido como o agente de *dispensou*. Essa característica da interpretação das coordenadas se repete em vários outros pontos e pode ser descrita informalmente como uma exigência de paralelismo sintático e semântico entre as duas orações. No caso de (36), o paralelismo se manifesta em que ambas as orações têm o mesmo agente (voltarei a este ponto mais adiante).

11.2.5.2. Orações sem NdP ou predicado

O aspecto mais importante do paralelismo que se observa entre estruturas coordenadas pode ser expresso no

Princípio de paralelismo das coordenadas: estruturas coordenadas pertencem sempre à mesma classe.

Por isso não é possível coordenar um verbo (finito) e um substantivo (* *correu e cansaço*), nem uma preposição e um verbo (* *chegou de e adorou Araxá*) etc. Assim, quando uma oração é coordenada a outra estrutura qualquer, entende-se que essa estrutura é outra oração. Em um caso como (35), isso não apresenta problema, porque os elementos coordenados apresentam ambos estrutura interna de oração. Mas há casos menos evidentes, como o de

(37) Serafim toca flautim e Ivone trombone.

A seqüência *Ivone trombone* não tem a estrutura interna de uma oração (na verdade, não tem a estrutura de nenhum tipo de constituinte da língua). Mas como está coordenada a uma oração (*Serafim toca flautim*), os falantes do português a tratam como uma oração, embora de certo modo defectiva. Em particular, *Ivone trombone* deverá ser semanticamente interpretada da maneira como o são as orações.

Ora, se isso acontecer, teremos na segunda oração dois SNs que não poderão receber papel semântico: o papel semântico depende do verbo da oração (ver R1 a R3, que se aplicam somente a frases com certos verbos), e essa oração não tem verbo. A estrutura não poderia receber uma interpretação bem formada e seria inaceitável, em virtude de um princípio geral

que **proíbe a presença de SNs sem papel temático**.

Como isso não acontece e (37) é aceitável, é necessário procurar uma regra que supra a ausência do verbo, pelo menos para efeitos semânticos. Essa regra deverá basear-se na presença de outra oração, esta completa, coordenada a *Ivone trombone*. Tanto é assim que, se essa outra oração não estiver presente, a seqüência será realmente mal formada:

(38) * Ivone trombone.

Intuitivamente, o falante percebe que (38) é mal formada porque "não se sabe o que Ivone está fazendo com o trombone" — ou seja, não há como determinar os papéis semânticos desses dois SNs.

A regra que resolve o problema é a seguinte:

Regra 11 (R11)

Inserir na representação semântica de uma oração coordenada sem NdP ou predicado uma cópia da matriz semântica do NdP ou predicado (respectivamente) da oração imediatamente precedente.

R11 introduzirá, então, na representação semântica de *Ivone trombone* uma cópia da matriz semântica de *toca*. Desse modo, *Ivone* poderá receber o papel de agente, e *trombone* o de paciente, resultando em uma interpretação bem formada.

Se a oração elíptica ocorrer sozinha, como em (38), R11 não se poderá aplicar, por falta de oração precedente, e o resultado será inaceitável, como vimos. E, como R11 requer a existência de uma oração precedente, também será inaceitável uma frase em que *Ivone trombone* venha em primeiro lugar:

(39) * Ivone trombone e Serafim toca flautim.

11.2.5.3. Orações sem OD e orações sem AC

Além dos casos vistos acima, ocorrem frases em que há a recuperação de outros termos elípticos da oração. Por exemplo, em

(40) Nair fez e Nadir vendeu os bolinhos.

entende-se que "Nair fez os bolinhos", embora a primeira oração não contenha o SN *os bolinhos*.

É claro que não seria possível simplesmente estender R11 para englobar também o objeto direto, porque R11 recupera um elemento elíptico da segunda oração e, em (40), a elipse está na primeira. Parece que teremos de formular uma nova regra. Vou incluir na formulação duas funções: objeto direto e adjunto circunstancial (AC); mas é quase certo que a regra vale também para outras funções. Convido o leitor a tentar estender a formulação a casos não cobertos pela regra.

Regra 12 (R12)

Inserir na representação semântica de uma oração coordenada sem OD ou AC uma cópia do OD ou AC (respectivamente) da oração imediatamente seguinte.

Esta regra é opcional.

No caso de (40), R12 inserirá a matriz semântica de *os bolinhos* na primeira oração, de maneira que entenderemos "os bolinhos" como paciente de *fez*. Note-se que, sem isso, a frase teria de ser inaceitável, porque o verbo *fazer* exige objeto direto e, portanto, a expressão de seu paciente.

Um exemplo com AC é o seguinte:

(41) Nair mostrou o aparelho e Nadir explicou seu funcionamento a Lair.

Em geral entende-se que Nair mostrou o aparelho a Lair.

Agora, por que R12 terá de ser opcional? Em (40), a sua não-aplicação resulta em inaceitabilidade; mas já vimos que isso é por causa da transitividade e das exigências semânticas de *fazer*. Já se a frase for

(42) Nair rouba e Nadir vende bolinhos.

será possível entender que (a) Nair rouba os bolinhos que Nadir vende, ou então que (b) Nair vive de roubar (algo não-especificado) e Nadir de vender bolinhos. Essa ambigüidade é possível porque *roubar*, ao contrário de *fazer*, não exige objeto (exigência sintática) nem paciente (exigência semântica). Assim, R12 é opcional; mas, como sempre, o resultado terá de ser submetido à inspeção das condições semânticas de boa formação. Essas condições inviabilizam a interpretação de (40) sem "os bolinhos" como paciente.

Note-se que, se R12 fosse obrigatória, a frase seguinte teria de ser inaceitável:

(43) Ronaldo ronca e Charles toca bumbo.

Se aplicarmos R12, a matriz semântica de *bumbo* será introduzida como paciente de *ronca*; mas esse verbo não admite paciente (nem objeto direto) e a representação semântica resultante será mal formada. Se R12 fosse obrigatória, essa seria a única interpretação possível de (43), que teria de ser inaceitável. O fato de (43) ser aceitável (em uma interpretação na qual "bumbo" não é paciente de "toca") mostra que R12 é opcional: a interpretação na qual R12 se aplica é mal formada, mas resta aquela que se obtém sem aplicar R12.

11.2.5.4. OD elíptico à direita

Além dos casos de elipse vistos acima, que encontram paralelo aproximado em muitas outras línguas, o português apresenta uma curiosa construção na qual o objeto direto da oração coordenada da direita é que falta e se interpreta em

função do OD da oração da esquerda. Um exemplo é a frase

(44) Fia fez a fantasia e Fernanda forrou.

Entende-se que "Fernanda forrou a fantasia". Observe-se como essa interpretação não pode ser derivada por R12, que só admite copiar a matriz semântica de um OD se este estiver em uma oração à direita, como em

(45) Fia fez e Fernanda forrou a fantasia.

Além disso, R12 não é exclusiva de objetos diretos, funcionando igualmente para o adjunto circunstancial e talvez para outras funções. Já o processo que estamos examinando só funciona com o objeto direto. Assim, se o testamos com um AC, como em

(46) Fia foi procurada por Fred, e Fernanda foi insultada.

não se entende que "Fernanda foi insultada por Fred"; o agente de *insultar* fica sem especificação e tanto pode ser Fred como outra pessoa qualquer.

O mesmo acontece com o atributo:

(47) Fia chegou furiosa, e Fernanda foi embora.

Não há nenhuma necessidade de se entender que "Fernanda foi embora furiosa".

Um terceiro fator que distingue R12 do processo ilustrado em (44) é a possibilidade de conservar um pronome na oração elíptica, sem nenhuma alteração semântica. Assim, (48) é bem formada e sinônima de (44):

(48) Fia fez a fantasia e Fernanda a forrou.

Já em casos como (45), onde é possível aplicar R12, a presença do pronome faz com que a frase fique inaceitável:

(49) * Fia fê-la, e Fernanda forrou a fantasia.

Como se vê, toda a evidência indica que temos aqui um processo diferente do que é ilustrado pelas frases às quais se aplica R12. É necessário, portanto, tratar (44) à parte.

Para isso, vou formular uma nova regra, a saber,

Regra 13 (R13)

Inserir na representação semântica de uma oração coordenada sem OD uma cópia da matriz semântica do OD da oração imediatamente precedente.

Essa regra se aplica da maneira que já deve ser familiar ao leitor e está sujeita às restrições semânticas usuais.

11.2.5.5. Elipses e pronomes

Vimos, no final da última seção, que certas frases admitem a presença de um pronome em vez da fal-

ta pura e simples do OD (elipse); nesses casos, não há nenhuma diferença de significado entre as duas versões da frase [o exemplo é o par (44)-(48)]. Isso nem sempre acontece com os casos de elipse em orações coordenadas. Por exemplo, vimos que (49) é inaceitável, justamente pela presença de um pronome:

(49) * Fia fê-la, e Fernanda forrou a fantasia.

Vamos agora examinar brevemente as condições que determinam a possibilidade de ocorrência de um pronome; veremos que, em certos casos [como em (44)-(48)], o pronome ou sua ausência alternam livremente; em outros, a presença do pronome causa inaceitabilidade [como em (49)]; e, finalmente, em outros casos, o pronome pode ocorrer, mas surge uma diferença de significado.

Vejamos então um exemplo deste último caso, pois os dois primeiros já foram exemplificados:

(50) Afonso chamou Careca e dispensou Newton.

Nesta frase, entende-se claramente que "Afonso dispensou Newton"; essa acepção nos é fornecida por R10 (ver 11.2.5.1.), sem maiores complicações. Mas se colocarmos o pronome sujeito em seu lugar,

(51) Afonso chamou Careca e ele dispensou Newton.

a interpretação mais imediata (e para alguns a única aceitável) será a que inclui "Careca dispensou Newton". Como se vê, o pronome explícito tem uma relação de correferência diferente da relação estabelecida para uma elipse.

Temos, portanto, três situações distintas para o pronome, ilustradas pelos exemplos (48), *(49) e (51). Temos aqui novamente uma situação de controle, tal como vimos em 11.2.4.2. para sentenças compostas por subordinação. E será necessário formular as condições em que um SN controla um pronome em estruturas coordenadas. As condições são as seguintes:

Definição de "SN controlador" para estruturas compostas por coordenação

SN1 controla opcionalmente *um pronome quando:*

SN1 precede o pronome;

SN1 é objeto direto;

SN1 e o pronome estão em orações diferentes e contíguas.

(A bem dizer, não fico muito feliz com essa formulação; darei algumas razões logo adiante.)

O controle neste caso é opcional, porque geralmente é possível entender o pronome como referindo-se a uma entidade não mencionada na frase. Assim, em (51), pode-se entender *ele* como referindo-se a *Careca* ou então a alguma outra pessoa não mencionada na sentença. Isso

explica por que a frase seguinte é aceitável:

(52) Afonso chamou Newton e ela dispensou Careca.

Se o controle fosse obrigatório, estabeleceria a correferência entre *Newton* e *ela*; e, como há choque de gêneros, o resultado deveria ser inaceitável. Sabemos que não é, porque *ela* não é obrigatoriamente controlado por *Newton* e pode referir-se a uma pessoa não mencionada no período.

Há problemas a investigar relativos a essa opcionalidade. Em certos casos, por razões obscuras, o controle parece ter "mais força". Assim, em

(48) Fia fez a fantasia e Fernanda a forrou.

é muito difícil entender o pronome como se referindo a outra coisa que não a fantasia (por exemplo, à camisa ou à calça).

Outro problema com o controle em coordenadas é que na frase

(53) César impressionava os homens e ele beliscava as mulheres.

muitos falantes admitem relacionar *ele* com *César*. Ora, como *César* não é objeto direto, não deveria ser um controlador possível neste caso. Pessoalmente, acho (53) de aceitabilidade marginal.

Não vou levar esta discussão mais adiante. Baste-nos, por ora, reconhecer que o princípio acima dá conta de boa parte dos dados, mas que certamente exige mais pesquisa, antes que se chegue a uma formulação satisfatória.

Exemplos como o do par (44)-(48) mostram que os pronomes se comportam diferentemente das elipses no que diz respeito a suas possibilidades de correferência com SNs presentes na frase.

Uma conseqüência importante disso é que a análise aqui adotada, que considera **sem sujeito** frases tradicionalmente consideradas como de sujeito oculto, deve ser mantida. Se a segunda oração de

(50) Afonso chamou Careca e dispensou Newton.

tivesse um "sujeito oculto", válido para efeitos de sintaxe, seria difícil explicar por que esse "sujeito oculto" se comporta, semanticamente, de maneira diferente da de um sujeito "claro" como *ele*: *ele* pode referir-se a *Careca* ou a uma pessoa não mencionada, ao passo que o "sujeito oculto" poderia referir-se exclusivamente a Afonso. Conclui-se que não há nenhum "sujeito (pronominal) oculto" na segunda oração de (50); esta, na verdade, não tem sujeito.

QUARTA PARTE: LEXICOLOGIA

12 Princípios de taxonomia

12.1. PARA QUE CLASSIFICAR AS PALAVRAS?

12.1.1. Classes e traços

Tomemos três palavras do português; digamos *cabeça, escola* e *sempre.* É possível mostrar que duas delas (*cabeça* e *escola*) têm muitas características gramaticais em comum, e que a terceira, *sempre,* é nitidamente diferente das outras duas.

Por exemplo, *cabeça* e *escola* admitem certos sufixos: *cabecinha, escolinha*; *cabeças, escolas.* Já *sempre* não admite esses sufixos (na verdade, não admite nenhum sufixo): * *semprinha*, * *sempres.* O acréscimo de sufixos (a formação da palavra) é do domínio da morfologia; assim, podemos dizer que *cabeça* e *escola* têm certos traços morfológicos em comum, a saber, a possibilidade de acrescentar as terminações *-inha* e *-s. Sempre* não tem esses traços e nisso se assemelha a outras palavras, como *hoje* e *porém.*

Por enquanto estamos lidando apenas com dois traços (o acréscimo de *-inha* e o de *-s*) e três itens léxicos, de modo que o exemplo é extremamente simplificado. Mas façamos de conta que esses poucos elementos nos autorizam a formular conclusões; eventualmente, como é óbvio, será necessário lidar com números muito maiores de traços e itens léxicos. Voltando aos nossos exemplos, poderíamos então dizer que as palavras *cabeça* e *escola* pertencem a uma **classe** diferente da que engloba a palavra *sempre.*

Antes de passar adiante, vamos refletir sobre o que significa, exatamente, colocar duas ou mais palavras na mesma classe.

A primeira vantagem de se definir classes é que se torna possível fazer afirmações gramaticais com o máximo de economia. Se não colocássemos *escola* e *cabeça* na mesma classe, teríamos de repetir a informa-

ção de que "aceita as terminações *-inha* e *-s*", separadamente para as duas palavras. Já, se as reunirmos na classe dos "substantivos", poderemos dizer, mais economicamente, que "os substantivos aceitam as terminações *-inha* e *-s*" (mas lembremo-nos que isto é apenas uma ilustração; essa afirmação não vale para todas as palavras usualmente chamadas de "substantivos").

A economia, no caso, pode parecer pequena: em vez de duas afirmações, podemos fazer uma. Mas as descrições gramaticais se ocupam de línguas inteiras, e aí a economia pode ser muito grande. Verificaremos que há outras palavras que se comportam da mesma maneira que *escola* e *cabeça*, como *mesa, página, orelha, cerveja, coelha, porta, força* etc. Há vários milhares de palavras como essas, que se comportam exatamente como *escola* e *cabeça* no que diz respeito à possibilidade de acrescentar aquelas terminações. Se as colocarmos todas na mesma classe, poderemos fazer nossa afirmação única no lugar de milhares de afirmações idênticas separadas.

Além disso, poderemos descobrir que há outras afirmações gramaticais que valem exatamente para essas mesmas palavras — ou seja, para essa mesma classe. Por exemplo, *escola* e *cabeça* podem ocorrer, sozinhas, como sujeito de uma oração:

(1) Escola às vezes é um aborrecimento.

(2) Cabeça é muito difícil de desenhar.

Ora, o mesmo se pode dizer de todas as outras palavras mencionadas:

(3) Mesa ocupa muito espaço.

(4) Coelha come mais do que coelho.

(5) Força não vai resolver essa questão.

e assim por diante. Havíamos visto que essas palavras tinham em comum certas características morfológicas: a possibilidade de serem acrescidas de certos sufixos. Agora descobrimos que também têm uma característica sintática em comum, ou seja, a possibilidade de formarem, por si sós, o sujeito de uma oração (a rigor, isso deve ser formulado assim: a possibilidade de formarem, por si sós, um sintagma nominal).

Estamos lidando, pois, não apenas com um grande número de palavras, mas também com um número maior de afirmações que valem para todas elas. Como se vê, em tais casos, o estabelecimento de classes se torna imperativo. Só assim fica possível fazer afirmações relativamente simples, como "os substantivos aceitam as terminações *-inha* e *-s*, e podem formar sozinhos um sujeito". À medida que descobrimos mais características que eles têm em comum, aumenta a utilidade da classe estabelecida. Essas características morfológicas, sintáticas ou semânticas que servem para classificar as palavras se denominam **traços**. Diremos que as palavras *escola, cabeça, força* etc. têm certos **traços** morfológicos e sintáticos em comum, o que nos possibilita colocá-las em uma só

classe. A classe (provisória!) dos "substantivos", que acabamos de estabelecer, se define através de três traços: são "substantivos" aquelas palavras que (a) podem receber a terminação *-inha*, (b) podem receber a terminação *-s* e (c) podem ser, sozinhas, sujeito de uma oração.

A bem dizer, a economia descritiva não é a única razão que nos leva ao estabelecimento de classes; por isso, disse eu acima que a economia é a "**primeira** vantagem de se definirem classes". Há outras razões, e importantes, mas aqui não será possível tratar delas; têm a ver com fatores tais como a depreensão dos grandes traços da estrutura da língua, ou o estudo da organização da memória para elementos lingüísticos. Na presente discussão, entretanto, vou enfatizar o aspecto da economia descritiva, que é o que nos interessa mais de perto em uma gramática.

12.1.2. Classificando por objetivos

No exemplo acima, utilizamos três traços, a saber: a possibilidade de acrescentar a terminação *-inha*; a de acrescentar a terminação *-s*; e a de constituir o sujeito de uma oração. Naturalmente, seria possível usar outros traços, que não esses, na classificação dessas palavras. Por que escolhemos esses traços, e não outros quaisquer?

Qualquer classificação (gramatical ou não) se faz em função de objetivos. Por exemplo, não há dúvida de que são grandes as diferenças entre um bebê e um homem de cinqüenta anos; segundo certos critérios, o bebê seria mais semelhante a um chipanzé recém-nascido: no tamanho, na forma do corpo, na incapacidade de falar etc. No entanto, para os objetivos da vida normal, não hesitamos de classificar juntos o bebê humano e o homem, em oposição ao chipanzé. Essa classificação se justifica porque, na imensa maioria dos casos, compensa tratar juntamente o bebê e o homem. Nenhum deles poderia ser colocado em exibição em um zoológico; ambos podem ser objeto de ensino escolar; ambos recebem nome e sobrenome (e, eventualmente, CPF). Nada disso se aplica ao chipanzé.

Mas agora imaginemos uma situação pouco comum: digamos que vamos construir assentos individuais para um avião que transportará homens adultos, chipanzés recém-nascidos e bebês. Agora o que se torna relevante é o tamanho, a forma e o peso do corpo; por conseguinte, para efeitos da confecção dos assentos, classificaremos de um lado os adultos e de outro (juntos) os bebês humanos e os chipanzés. Um novo objetivo levou, necessariamente, a uma nova classificação.

Assim, de início é importante enfatizar certos fatos acerca das classificações. Uma classificação, em si,

não tem muito interesse; ela interessa na medida em que serve a um objetivo. Em outras palavras: os critérios de classificação possíveis são infinitamente variados, e a seleção de critérios se faz em função dos **objetivos** da classificação. O mesmo conjunto de elementos será classificado diferentemente caso mudem os objetivos.

Assim, a primeira pergunta a ser formulada, quando se pensa em classificação, tem a ver com os objetivos imediatos dessa classificação. Se queremos elaborar uma classificação das palavras do português, muitas opções estarão abertas, como por exemplo:

(a) classificá-las por ordem alfabética;

(b) classificá-las conforme o número de letras de cada uma;

(c) classificá-las segundo sua origem.

A primeira dessas classificações faz sentido para alguém que esteja compilando um dicionário. Assim, é importante saber que *abacate* deve vir antes de *araticum*, e ambos antes de *banana*. A segunda classificação pode ser útil para uma pessoa que esteja planejando um *outdoor*. E a terceira é importante para o estudioso da história externa da língua.

É também possível imaginar classificações que não seriam úteis em nenhuma situação razoável. É o caso de classificações que seguem critérios heterogêneos e não-relacionados. Por exemplo, poderíamos definir uma classe composta das palavras que têm mais de doze letras e significado abstrato. O que há de absurdo nessa classe não está na sua formulação propriamente dita, pois esta é bastante clara (desde que se consiga conceituar claramente o que vem a ser "significado abstrato"). O absurdo está em que essa classe de palavras não teria utilidade em nenhuma situação descritiva. Os membros dessa classe não têm em comum nenhum traço gramatical relevante: a classe incluiria as palavras *estruturalismo* e *renascentismo*, mas não *gerativismo*, *darwinismo*, *renascentista* ou *estruturalista*. Fica bem evidente que essa classificação não exprime nada de útil, gramaticalmente falando; em outros termos, não se consegue ver um objetivo para ela que seja útil em gramática.

É importante fazer estas considerações porque algumas das classes encontradas nas gramáticas tradicionais parecem ser propostas sem a preocupação de explicitar os objetivos a que deveriam servir. Numa gramática tradicional, há muitos tipos de classes e algumas são úteis, ao passo que outras são mais ou menos gratuitas; conforme apontei acima, a divisão de elementos em classes não tem interesse, a menos que esteja vinculada a um objetivo. No nosso caso, esse interesse é o de descrever a estrutura da língua. Vejamos alguns exemplos.

A classe usualmente chamada dos "substantivos" costuma ser dividida em subclasses. Assim, os substantivos se distinguem em masculi-

nos e femininos, o que é necessário a fim de se descrever o fenômeno da concordância nominal: *relógio* **novo**, *pulseira* **nova**. Se não distinguíssemos masculinos de femininos, a única maneira de descrever esse fenômeno seria dar uma lista completa de todos os itens que aceitam *novo* e dos que aceitam *nova*. Desse modo, é preciso reconhecer que essa subclassificação dos substantivos faz algum sentido para a descrição da estrutura do português.

Por outro lado, não tem utilidade distinguir as palavras segundo o número de sílabas: monossílabos, dissílabos, trissílabos (tetrassílabos?), polissílabos. Essas classes não são utilizadas na descrição, e seria melhor passar sem elas. O mesmo vale, pelo menos no que diz respeito à morfossintaxe, para a distinção entre substantivos concretos e abstratos. Em inglês, os dois grupos têm comportamento sintático distinto; os abstratos não levam artigo, como em

(6) Linguistics is a science.

"A lingüística é uma ciência."

(7) * The linguistics is a science.

Mas em português isso não acontece: os substantivos abstratos têm comportamento formal idêntico ao dos concretos. A se confirmar essa asserção, a distinção concreto/abstrato deverá ser abandonada no estudo da morfossintaxe da língua (a distinção pode ser útil na semântica, claro).

Exemplos como esses, de classificações seriamente inadequadas para a descrição da estrutura da língua, são numerosos. O problema provém em parte da atitude dos gramáticos, que não se preocupam em justificar previamente as classificações propostas, contentando-se em repetir o que a tradição fornece. Em outras palavras, falta consciência dos objetivos da classificação. A situação é tal que se torna necessário refazer a maior parte do trabalho de classificação das palavras: poucas das classes tradicionais podem ser consideradas adequadas. Por conseguinte, neste setor da gramática, talvez mais do que em qualquer outro, será inevitável propor mudanças radicais.

O trabalho, como se verá, mal está começando. Relativamente poucas classes já foram estudadas; não é possível dar no momento uma lis-ta exaustiva das classes de palavras do português. Pode-se, no máximo, apresentar as classes já depreendi-das e esboçar algumas outras cuja depreensão é provável. Se isso parecer ao leitor um passo atrás em relação ao que oferecem as gramáticas atuais, convido-o a submeter qualquer das classes tradicionais a uma crítica; formule perguntas como as seguintes:

— Que relação existe entre a definição da classe e o conjunto de palavras habitualmente consideradas como pertencentes a ela? A se aplicar a definição de maneira estrita, chegar-se-á àquele mesmo conjunto de palavras?

— A definição, em si, é suficientemente clara e aplicável ou, ao contrário, lança mão de noções desconhecidas, vagas e misteriosas?

— Qual é a utilidade da classe proposta para a descrição da língua? Ou seja, a classe funciona como um conjunto mais ou menos coeso, para efeitos do funcionamento de partes importantes da gramática?

Creio que o leitor encontrará um número surpreendente de classes que deixam de atender a qualquer critério coerente, tanto em termos de sua formulação quanto em termos de sua utilidade para a descrição da estrutura. Quanto às definições, não sei de uma única que realmente funcione. E das classificações feitas na prática algumas poucas são razoavelmente claras (mas nunca decorrem das definições; decorrem antes do que chamei "doutrina gramatical implícita"; ver Perini, 1985, p. 15); por exemplo, a classe dos verbos. A maioria, porém, são classes incoerentes (pronomes, advérbios) ou de limites difusos, mas não reconhecidos como tais (adjetivos, substantivos).

Se esta *Gramática* é menos informativa a respeito das classes do português do que as gramáticas tradicionais, isso se deve à recusa em apresentar classes que não sejam devidamente fundamentadas em princípios tais como: aplicabilidade da definição e sua adequação ao conjunto de palavras incluído na classe; coerência da classe, isto é, relativa uniformidade dos membros quanto ao comportamento gramatical; e importância da classe na descrição dos grandes traços da língua. Essa tomada de posição nos força a incluir na gramática as limitações do conhecimento gramatical atual — em particular, os resultados relativamente parcos da pesquisa na área de classificação das palavras. Há, entretanto, alguns projetos em andamento nessa área, e em breve se tornará possível dar um panorama muito mais completo das classes de palavras do português.

12.2. COMO CLASSIFICAR?

12.2.1. O formal e o semântico

Em termos muito gerais, pode-se dizer que a classificação das palavras (e das unidades morfossintáticas em geral) tem como objetivo agrupar aquelas que têm **comportamento gramatical semelhante**. Isso nem sempre se coloca nesses termos, mas seria aceitável para a maioria dos gramáticos, qualquer que seja sua convicção teórica. O ponto controverso é o que se deve entender por "comportamento gramatical"; e a questão mais crucial aqui é a de como colocar a oposição entre a descrição das formas e a do significado.

Como tenho afirmado repetidamente neste livro, acredito que a posição correta a esse respeito é a de uma separação estrita entre a descrição dos aspectos formais (morfossintáticos) e os semânticos. Esse pressuposto condiciona toda a análise: segundo a posição que se tome frente ao problema, variam os próprios objetivos da descrição lingüística. Se se aceita que um dos objetivos principais da lingüística é explicitar a relação que existe entre a forma e o significado, a separação metodológica dos dois aspectos se impõe. Essa me parece ser a posição de Saussure e também a de Chomsky; ver, a respeito, Perini, 1985, p. 21-27, e a seção 2.1.2. desta *Gramática*. Nessas passagens e em outras, explicito minha crença de que o estudo da relação forma/significado depende de um estudo separado de cada um desses aspectos da linguagem. Não se trata, evidentemente, de uma posição pessoal, isolada, e muito menos de inovação minha; a maioria dos lingüistas em atividade hoje admite a validade desse princípio, que, aliás, já estava bem estabelecido há pelo menos uns cinqüenta anos.

Uma vez decidido que vamos estudar separadamente os dois aspectos, essa separação se transfere automaticamente às classes de formas e, em particular, às classes de palavras. Assim, é necessário classificar as palavras quanto a seus traços formais, isto é, quanto ao seu comportamento sintático e morfológico; e também é necessário classificá-las quanto a seus traços de significado. Mas não se espere encontrar classes idênticas, ou mesmo semelhantes, quando se aplicarem os dois conjuntos de critérios; duas palavras que se reúnem claramente do ponto de vista semântico podem ter comportamento sintático muito diverso, e vice-versa. No final, o resultado deverá ser não uma classificação única das palavras, mas pelo menos duas classificações distintas: uma formal, a outra semântica.

Vejamos um exemplo, para terminar esta seção; mais adiante, aparecerão outros exemplos, que ajudarão a deixar mais clara a natureza do problema.

As palavras *sim* e *não* são geralmente classificadas juntamente como "advérbios". Do ponto de vista semântico, talvez haja alguma razão nisso: as duas palavras exprimem noções relacionadas, pelo menos em situações como as seguintes:

(8) O táxi está livre? — Sim.

(9) O táxi está livre? — Não.

Nesses casos, *sim* e *não* servem para atribuir um valor de verdade à frase interrogativa. Ou seja, em (8) a resposta pode ser parafraseada como "a frase *o táxi está livre* é verdadeira"; e em (9) teríamos "a frase *o táxi está livre* é falsa". Como se vê, há pelo menos alguma razão para juntar essas palavras em uma classificação de base semântica.

Mas as diferenças sintáticas entre *sim* e *não* são muito profundas, a ponto de termos de reconhecer que não faz sentido classificá-las juntamente do ponto de vista formal. Casos como (8) e (9), onde *sim* e *não* ocorrem em ambientes sintaticamente idênticos, são a exceção. Em geral, suas condições de ocorrência na frase são diferentes, como, por exemplo, em

(10) a. O táxi *não* está livre.
b. * O táxi *sim* está livre.

(11) a. Os *não*-convidados levarão um crachá especial.
b. * Os *sim* convidados levarão um crachá especial.

Em geral, parece que *não* pertence a uma pequena classe, que desempenha, entre outras, a função de negação verbal. Já *sim* tem semelhanças de comportamento com as palavras geralmente chamadas "interjeições", e valeria a pena apurar se não se trata de uma delas. De qualquer maneira, é muito claro que *sim* e *não* têm comportamento sintático bem diferente e que, portanto, não devem ser colocadas na mesma classe, no que pese sua semelhança de conteúdo semântico.

Essa situação se repete com freqüência. Por outro lado, certamente há alguma relação entre as classes de palavras (estabelecidas formalmente) e certas categorias semânticas. Apenas, isso não é um ponto de partida para a análise, mas antes uma questão a investigar. Um dos aspectos da relação som/significado, cuja investigação é um dos objetivos da lingüística, é justamente o grau de coincidência que se verifica entre as classificações feitas com base na forma e as feitas com base na semântica. É exatamente por isso que se insiste tanto na necessidade de estudar separadamente esses dois aspectos da linguagem: de outro modo, o estudo da relação entre eles estaria viciado desde o início.

12.2.2. Traços: primeira e segunda ordem

Acabamos de ver que classificar as palavras implica em elaborar uma classificação sobre critérios formais (sem excluir da descrição a classificação semântica, mas separando-se nitidamente dela). Vejamos agora como estabelecer essa classificação formal (morfossintática): que critérios utilizar e como representá-los na gramática.

A noção fundamental a ser utilizada aqui é a de **potencial funcional.** Entende-se por potencial funcional o conjunto de funções sintáticas que uma palavra pode desempenhar.

Assim, a palavra *corríamos* pode ser núcleo do predicado, e apenas isso; já a palavra *Maria* pode ser núcleo de um SN, mas não núcleo do predicado, nem modificador, nem negação verbal. A palavra *velho* pode

ser núcleo de um SN, modificador externo ou pré-núcleo. Assim, essas três palavras diferem em seu potencial funcional. Uma maneira de indicar isso é definir uma série de traços derivados das funções; diremos, então, que *corríamos* é marcado [+NdP] e, naturalmente, também [–NSN, –NV, –ModE] etc. (Para simplificar a notação, convencionemos que todos os traços que não aparecem marcados positivamente são automaticamente entendidos como tendo marca negativa.)

Traços como esses são de natureza diferente dos que utilizamos na definição das funções. Os traços [CV], [Cl], [Ant] etc., apresentados na seção 3.2.1., referem-se diretamente a propriedades sintáticas dos termos analisados. Já um traço como [+ModE] é a abreviatura de um feixe de traços: um elemento marcado [+ModE] é um elemento que pode ocorrer no último lugar no SN máximo, não pode ocorrer no primeiro lugar etc. Um elemento marcado [+OD] é um elemento que pode ocorrer marcado como [–CV, +Ant, +Q, +Cl, –PA, –pNdP]. Assim como as funções propriamente ditas são abreviaturas de matrizes de traços, traços como [OD] são abreviaturas da propriedade de ocorrer marcado com determinada matriz. A estes traços chamaremos **de segunda ordem**, em oposição aos traços já vistos, como [Ant], [Cl], [CV] etc., que são de **primeira ordem**.

O potencial funcional de um item se exprime através de traços de segunda ordem. Voltando aos exemplos acima, podemos exprimir o potencial funcional das palavras vistas da seguinte maneira:

corríamos: pot. func. [+NdP];

Maria: pot. func. [+NSN];

velho: pot. func. [+NSN, +ModE, +PN].

Cada função sintática, seja de nível oracional, seja de nível suboracional, serve de base para a formulação de um traço de segunda ordem.

Como primeira aproximação, então, podemos dizer que **pertencem a uma mesma classe palavras que têm potencial funcional idêntico**; ou, para dizer a mesma coisa em outros termos, pertencem a uma mesma classe palavras marcadas com a mesma matriz de traços de segunda ordem.

Segundo essa definição, então, a palavra *trabalhavam* deve ficar na mesma classe que *corríamos*, porque também é marcada [+NdP]; e a palavra *amigo* fica na mesma classe que *velho*, pois é marcada [+NSN, +ModE, +PN]. É assim (como primeira aproximação, repito) que se definem as classes de palavras em sintaxe. Há mais a dizer a respeito, porque os fatos não são nada simples. Mas vale a pena determo-nos por algum tempo nessa conceituação simples, examinando suas conseqüências para a análise e para a pesquisa.

Observe-se, finalmente, que alguns traços de segunda ordem podem ser definidos não precisamente através de uma função, mas através da coocorrência com um termo de determinada função. Assim, utilizaremos mais adiante o traço [Int–], que significa **a propriedade de ocorrer em um sintagma juntamente com um intensificador**. Esses traços (**traços de contexto sintático**) são também considerados de segunda ordem e são tratados da mesma maneira que os traços que exprimem estritamente o potencial funcional.

12.3. CLASSES E FUNÇÕES

Classes são entidades gramaticais nitidamente distintas de **funções**; no entanto, encontra-se freqüentemente alguma confusão entre elas. Por isso, parece-me necessário fazer aqui um esclarecimento (retomando um tema já abordado em Perini, 1985 e 1989).

A **função** é um princípio da organização da oração; determinar a função de um constituinte é formular sua relação com os demais constituintes da unidade de que ambos fazem parte. Assim, por exemplo, diremos que *Pedro* é o sujeito da oração

(12) Pedro almoçou comigo.

Isso equivale a dizer que *Pedro* está em relação de concordância com *almoçou*. Já em

(13) Eu convidei Pedro para o almoço.

Pedro não é sujeito porque não apresenta aquela relação com *almoçou*. Como se vê, "ser sujeito" é simplesmente ter uma relação particular com os outros elementos da oração [poderíamos acrescentar que *Pedro* em (12) vem em primeiro lugar, que não poderia ser posposto etc.].

Já a **classe** é uma propriedade que se atribui a um elemento fora de contexto. Voltemos aos exemplos (12) e (13): uma coisa que esses exemplos nos dizem é que *Pedro* **pode ser** sujeito, como em (12), ou objeto direto, como em (13). Essa afirmação vale para o item *Pedro*, independentemente da oração em que ocorrer. Em (12), *Pedro* é sujeito — mas o item *Pedro* que ocorre como sujeito em (12) **pode ser** objeto direto; e é essa potencialidade que o classifica. Dizemos, simplificando um pouco as coisas, que *Pedro* é um substantivo (tipo 2; ver o capítulo 13) porque **pode ser** sujeito ou objeto direto. Esse é, como sabemos, o seu potencial funcional.

Em outras palavras, um elemento lingüístico **está desempenhando** uma função quando ocorre em determinada oração; e **pertence** a uma classe, independentemente da oração em que ocorrer. Por isso não faz sentido perguntar (sem citar uma oração) qual é a função de *Pe-*

dro; só se pode perguntar qual é sua função na oração tal. Mas pode-se perfeitamente perguntar qual é a classe a que *Pedro* pertence: trata-se de um substantivo, seja qual for a oração em que ocorrer. É claro que a classificação de um item é depreendida a partir da observação do comportamento desse item dentro das estruturas; mas é formulada de maneira geral, válida para o item em qualquer contexto. As palavras, especificamente, são classificadas em seu "estado de dicionário"; mas só adquirem função a partir de sua inserção em uma estrutura bem formada.

Uma decorrência importante dessa concepção de classe é que não se pode dizer que uma palavra pertence a uma classe em determinada oração, mas a outra classe em outra oração; ou, o que é o mesmo, que "funciona como membro de outra classe" em certas construções. Por exemplo, é incoerente dizer que *amigo* é um substantivo em (14), mas um adjetivo em (15):

(14) Meu amigo me traiu.

(15) Preciso de uma pessoa amiga.

Tanto em (14) quanto em (15), *amigo* pertence à mesma classe (é adjetivo, como veremos no capítulo 13). Sua classe se define, em parte, por poder ser núcleo de um SN — como em (14) — ou modificador externo — como em (15).

12.4. CLASSES FECHADAS E CLASSES ABERTAS

Antes de encerrar o capítulo, vou fazer uma distinção de importância metodológica e teórica: a distinção entre **classes abertas** e **classes fechadas.** Vou conceituar esses dois tipos de classes tomando como exemplo a análise do sintagma nominal que foi desenvolvida no capítulo 4.

Consideremos, por exemplo, as funções denominadas **determinante**, **possessivo**, **reforço**, **quantificador**, **pré-núcleo interno** e **numerador**. Cada uma dessas funções é desempenhada por um grupo pequeno e limitado de itens léxicos, ou seja, por uma **classe fechada** de itens. Assim, os itens que podem ser Det em português são menos de uma dezena: *o, um* (artigo), *este, esse, aquele,* e talvez *cada, todo* (no singular), *que* (todos com seus femininos e plurais, quando os possuem). Os itens que podem ser Poss são em número igualmente restrito: *meu, seu, nosso.* E os que podem ser Qf são em número igualmente limitado (se considerarmos em bloco a lista dos numerais ordinais): *poucos, muitos, único, vários, diversos,* mais os numerais ordinais.

Não apenas essas listas são breves, mas são comuns a todos os falantes da língua: não há falante do português que não conheça todos esses itens. Aprendemo-los nos primeiros

estágios da aquisição da língua e nunca mais os esquecemos.

Já as funções de **pré-núcleo externo**, **núcleo do SN**, **modificador interno** e **modificador externo** são desempenhadas por **classes abertas**. A lista dos itens que podem ser PNE, ou NSN, ModI ou ModE é muito grande: milhares ou mesmo dezenas de milhares de palavras; e ainda, no caso dos modificadores, um número indeterminado de sintagmas, como *de Salvador, de confiança* etc. Além disso, os falantes nunca têm um conhecimento completo dessas listas, e estão sempre aprendendo novos elementos, e esquecendo outros. Tome-se, como exemplo, a lista dos possíveis núcleos de um sintagma nominal; assim como se pode dizer

(16) Meu computador

onde *computador* é o NSN, pode-se também dizer

(17) a. Meu livro

b. Meu pai

c. Meu relógio

d. Minha gata

etc.

É fácil convencer-se de que a lista é imensa. E, embora qualquer falante do português conheça as palavras *computador, livro, pai, relógio* e *gata*, há muitos possíveis núcleos do SN que só são conhecidos por alguns falantes:

(18) a. Meu incunábulo

b. Minha tendinite

c. Meu clavicórdio

etc.

Por isso é que o conjunto das formas que podem ser NSN é chamado uma **classe aberta**.

As classes de palavras (como as classes de formas em geral) de uma língua se colocam, portanto, dentro de uma dessas duas categorias; em alguns trabalhos tradicionais encontra-se essa oposição, quando se distinguem "semantemas" (elementos de classe aberta) e "morfemas" (elementos de classe fechada). Prefiro, contudo, não adotar essa terminologia, porque é semanticamente motivada e nem todo elemento de classe fechada tem um significado puramente gramatical, estrutural; vejam-se, por exemplo, os possessivos, cujo significado é bastante concreto.

13 Classes de palavras em português

Conforme vimos no capítulo anterior, não será possível dar aqui uma lista completa das classes de palavras em português, pois as classes não foram ainda devidamente pesquisadas. A lista de classes habitualmente apresentada nas gramáticas pretende ser exaustiva, ou quase (se excluirmos as "palavras de classificação duvidosa" ou "de classificação à parte"). Mas tal se deve principalmente ao fato de que as classes tradicionais não são estabelecidas segundo critérios de coerência e relevância gramatical.

Na verdade, a classificação tradicional tem pouca utilidade para a descrição. Algumas classes são razoavelmente bem estabelecidas, mas nenhuma é bem definida; e as definições não costumam ter nada a ver com as classes propriamente ditas. A maioria das classes tradicionais simplesmente não fazem sentido em termos de descrição da língua.

13.1. VERBOS

Uma classe tradicional que se pode considerar bem estabelecida é a dos **verbos**. Sabemos que os verbos têm realmente um comportamento morfossintático muito homogêneo: flexionam-se da mesma maneira e desempenham a mesma função sintática. Por outro lado, a definição de "verbo" deixa muito a desejar; Cunha & Cintra, definem o verbo como

> [...] uma palavra de forma variável que exprime o que se passa, isto é, um acontecimento representado no tempo. [1985, p. 367]

Essa definição já foi criticada em outro trabalho (Perini, 1985, p. 23-27). Aqui basta observar que seria bem difícil aplicá-la a casos concretos. Diante de uma palavra de classificação desconhecida, como decidir (com base na definição) se é um verbo? Que quer dizer "exprime o que se passa"? Que vem a ser "um aconte-

cimento representado no tempo"? Não se pode dizer que em

(1) Estou impressionado com o sucesso de Pedrinho.

o sucesso de Pedrinho exprime algo que se passa? Mas onde está o verbo em *o sucesso de Pedrinho*?

E que significa "representado no tempo"? Em uma frase como

(2) A água se compõe de hidrogênio e oxigênio.

pode-se argumentar que o acontecimento não se representa no tempo, pois a afirmação vale para todos os tempos, e não para alguma ocasião especial. No entanto, há um verbo em (2).

O que há de claro e característico nos verbos — e que nos permite identificar os verbos sem grande dificuldade — são seus traços morfossintáticos. Um verbo é um lexema que varia em número, pessoa e tempo. Ou, para nos expressarmos em termos mais técnicos,

Verbo é a palavra que pertence a um lexema cujos membros se opõem quanto a número, pessoa e tempo.

(Chamamos **lexema** um conjunto de palavras que se distinguem através de flexão; assim, *corro, correr, corríamos* fazem parte de um lexema; *casa, casas* fazem parte de outro. Mas *casa* e *casebre* não fazem parte do mesmo lexema, porque se distinguem por derivação, e não por flexão.)

Em geral, as pessoas com algum treinamento gramatical não têm problemas em identificar os verbos. Ao considerarem uma palavra como *respondessem*, comparam-na com outros membros de seu lexema (seu paradigma) e verificam que estes se opõem quanto a pessoa: *respondesse/respondessem/respondêssemos; quanto a tempo: respondessem/respondam/respondem/responderão*; e quanto a número: *respondessem/respondesse.* Nisso é que nos baseamos para reconhecer os verbos — não em suas propriedades de exprimir acontecimentos representados no tempo; nem, para citar outra definição comum, em suas propriedades de exprimir ações, estados ou fenômenos.

Em outras palavras, a noção corrente de "verbo" é formal, e não semântica. O conjunto das palavras que exprimem uma ação, por exemplo, não se chama em geral de "verbos": *correr, corria, corrida, vingança, vingar, traiu, traição* etc. Como se vê, todas essas palavras exprimem, de alguma forma, uma ação; mas nem todas são verbos.

A classe dos verbos é uma das poucas que se pode definir e estabelecer no momento com clareza. A definição dada acima baseia-se nos traços morfológicos da palavra (variação em pessoa, tempo e número). Podemos acrescentar que o verbo é a única palavra que pode desempenhar a função sintática de núcleo do predicado.

Este último traço, entretanto, não pode ser incluído na definição de verbo, porque a noção de "núcleo do predicado" depende, ela própria, da identificação prévia dos verbos. Assim, vamos dar o primeiro passo na classificação das palavras do português da seguinte maneira:

(a) verbos são palavras que variam em pessoa, tempo e número;

(b) somente os verbos podem desempenhar a função de NdP.

Em (a), temos a definição (informal) de verbo; em (b), uma informação extra sobre o comportamento gramatical dos verbos (informação essa que não faz mais que repetir a postulação de NdP).

Como se vê, no caso do verbo, a definição é simples; e a descrição do comportamento gramatical da palavra é igualmente simples. A maior parte das classes, entretanto, apresenta complexidade muito maior de comportamento. Por isso mesmo, geralmente é muito mais difícil identificá-las; a maioria das pessoas acha mais difícil distinguir um adjetivo de um substantivo, ou um advérbio de uma conjunção, do que um verbo de qualquer dessas classes.

Em (b), acima, temos uma expressão do potencial funcional dos verbos: eles podem ser núcleos do predicado e nada mais. Essa deveria ser a definição de verbo; mas o verbo, justamente por ser o ponto de partida de nossa análise da oração, acaba sendo tratado diferentemente de todas as outras classes.

A definição que adotamos [resumida em (a)] é, na verdade, também a expressão de um potencial funcional — mas não da palavra a que chamamos "verbo"; é o potencial funcional de um radical, definido em nível morfológico. Assim, (a) pode ser traduzido mais ou menos como "verbo é toda palavra cujo radical pode coocorrer com os sufixos de modo-tempo e pessoa-número". Com essa única exceção, porém, as classes de palavras serão sempre definidas pelo seu potencial funcional em nível sintático.

13.2. SUBSTANTIVOS E ADJETIVOS

Ao contrário do que se dá com os verbos, as classes tradicionalmente denominadas "substantivo" e "adjetivo" têm limites muito pouco claros. É fácil distinguir formalmente um substantivo de um verbo, ou um adjetivo de um verbo; mas a separação entre substantivos e adjetivos é tão pouco marcada que há razões para duvidar da existência de duas classes distintas. Nesta seção, vou explorar essa separação; ao mesmo tempo, examinarei o problema central que dificulta o trabalho de classificação das palavras: a falta de distinção nítida entre os diversos tipos de comportamento gramatical. Veremos também que existe uma maneira de descrever com rigor esses tipos de comportamento gramatical e que

essa maneira nos leva a uma classificação significativamente diferente da tradicional.

Um traço (de segunda ordem) que nos pode servir para caracterizar os substantivos é a possibilidade de ocorrer na função de núcleo de um SN: [+NSN]. Essa é certamente a função mais típica dos "substantivos" da gramática tradicional. Em uma frase como

(3) Patrícia comprou um PC.

Patrícia é núcleo de um SN (na verdade, constitui, por si só, um SN). Essa é a posição mais "normal" de um substantivo. Portanto, vamos estabelecer o traço [+NSN] como parte da definição de uma classe que, tentativamente, identificaríamos com a dos substantivos.

Entre as palavras cujo potencial funcional inclui [+NSN], podemos citar: *Patrícia, inimigo, cabelo, fazendeiro.*

Todas essas palavras são tradicionalmente chamadas "substantivos". No entanto, encontramos também como NSNs as palavras *ela, endividado, verde, vencer, velho,* que não são "substantivos" segundo a taxonomia tradicional. Por exemplo,

(4) Ela comprou um PC.

(5) Os endividados vão ter um alívio com a nova lei.

(6) O verde está muito na moda.

(7) Vencer é meu objetivo.

(8) Velho em geral dorme pouco.

No caso de *ela*, pode-se dizer que se trata, de certa forma, de um substantivo (um "pronome substantivo"), de modo que não há necessariamente problema. Mas as outras palavras não costumam ser chamadas de substantivos e, no entanto, são inegavelmente [+NSN]. Seguindo este nosso primeiro critério, todas essas palavras deveriam ser colocadas na mesma classe.

Mas temos aí apenas um traço; e o potencial funcional das palavras, na maioria das vezes, se exprime por vários traços — o que equivale a dizer que a maioria das palavras pode desempenhar mais de uma função sintática.

Algumas das palavras examinadas podem ser modificadores, como em

(9) Uma mesa *velha*

(10) Um avião *inimigo*

(11) Um rapaz *endividado*

etc. Examinando o conjunto delas, verificaremos que o traço [+Mod] pode ser atribuído a *inimigo, fazendeiro, endividado, verde, velho.*

Por outro lado, são marcadas [–Mod], por não poderem ser modificadores, as palavras *Patrícia, cabelo, ela, vencer.*

Aqui já temos um corte que não corresponde a nenhuma das distinções da gramática tradicional. Seria de esperar que o traço [+Mod] servisse para caracterizar os "adjetivos" (a função de modificador seria a mais típica dos adjetivos); mas pa-

lavras como *fazendeiro* não são nunca classificadas como adjetivos, embora possam ter a função de ModE:

(12) Tenho um sobrinho fazendeiro.

Já as palavras que podem ser modificadores, mas não NSNs, são sempre chamadas "adjetivos": *exato, presidencial, materno.*

Resumindo a situação, temos até agora três tipos de comportamento gramatical, que definiriam, em princípio, três grupos, que presumivelmente poderão vir a formar classes distintas:

[+NSN, +Mod]:
inimigo, fazendeiro, endividado, verde, velho;

[+NSN, –Mod]:
Patrícia, cabelo, ela, vencer;

[–NSN, +Mod]:
exato, presidencial, materno.

O leitor poderá ter notado que não estou distinguindo aqui o modificador externo do interno. É que ainda não disponho de levantamentos léxicos feitos com base nessa diferenciação de funções. Assim, estou adotando o traço provisório [Mod], entendido como "a propriedade de ser modificador interno ou externo". Embora deixando escapar, com certeza, alguns detalhes importantes, esse traço será suficiente para a classificação preliminar que desenvolverei nesta seção.

A classificação dada acima não corresponde à tradicional de nenhuma forma. Não seria solução aceitarmos a idéia corrente de que um adjetivo pode ser "substantivado", ou vice-versa, pois essa é uma noção insustentável, conforme mostrei em outro lugar (ver Perini, 1985, p. 83, e os comentários do capítulo 12).

A observação principal a ser feita é a seguinte: uma palavra como, digamos, *inimigo* não é um substantivo que às vezes se "transfere" para a classe dos adjetivos, nem o oposto. Trata-se, pura e simplesmente, de uma palavra cujo potencial funcional inclui tanto a possibilidade de ser núcleo de um SN quanto a de ser modificador. Não há razão para se considerar uma dessas funções como básica, e a outra como derivada, já que ambos os sintagmas abaixo são perfeitamente normais:

(10) Um avião inimigo [*inimigo*: Mod]

(13) Um inimigo terrível [*inimigo*: NSN]

Vejamos agora um terceiro traço, definido como a possibilidade de ocorrer como pré-núcleo: [+PN] (novamente me vejo forçado a não considerar a distinção entre pré-núcleo interno e externo, pelas mesmas razões dadas acima para a não-distinção entre modificador interno e externo). Quatro das palavras vistas podem ser pré-núcleos:

(14) Um endividado comerciante

(15) Verdes mares

(16) O velho presidente

(17) O exato momento

Essas palavras serão marcadas [+PN], e as outras [–PN]. A divisão estabelecida por esse novo traço não coincide com as divisões já feitas com base nos outros dois traços; nem sequer corta um dos grupos já estabelecidos em dois subgrupos. Em vez disso, os itens marcados [+PN] podem pertencer ao segundo ou ao terceiro dos grupos anteriormente definidos. Como resultado, acabamos ficando com cinco grupos, a saber,

1º: [+NSN, –Mod, –PN]
Patrícia, cabelo, ela, vencer;

2º: [+NSN, +Mod, –PN]
inimigo, fazendeiro;

3º: [+NSN, +Mod, +PN]
endividado, verde, velho;

4º: [–NSN, +Mod, –PN]
presidencial, materno;

5º: [–NSN, +Mod, +PN]
exato.

Esses cinco grupos se distribuem em uma espécie de gradação: o primeiro grupo compreenderia os itens de comportamento mais tipicamente "substantivo"; o quarto e o quinto grupos, os itens mais tipicamente "adjetivos". O segundo e o terceiro grupos incluiriam itens de comportamento ambíguo: "substantivos", por poderem ser núcleos de SNs; mas também "adjetivos", por poderem ser modificadores e/ou pré-núcleos.

Mas essa interpretação peca por dar importância excessiva às noções de "adjetivo" e "substantivo". Afinal, que vem a ser exatamente um "substantivo"? Se definirmos "substantivo" como o item marcado [+NSN], estaremos incluindo nessa classe tanto os itens do primeiro grupo quanto os do segundo e do terceiro — e, portanto, a classe dos "substantivos" será heterogênea: todos os itens podem ser NSN, mas alguns podem ser também Mod, outros Mod ou PN, outros nem Mod nem PN. Em outras palavras, essa definição de "substantivo" será pouco útil, por deixar de exprimir uma parte importante do potencial funcional dos itens classificados; seremos, provavelmente, forçados a distinguir subclasses dentro da classe maior.

Já se definirmos "substantivo" como o item marcado [+NSN, –Mod, –PN], e "adjetivo" como o item marcado [–NSN], os itens do segundo, terceiro e quarto grupos terão de ser incluídos em outras classes. Ou seja, duas classes apenas, de qualquer forma que se definam, não são suficientes para descrever o comportamento gramatical desse conjunto de itens.

Vou deter-me um pouco nesse ponto, importante por ir contra uma tradição muito antiga e arraigada. Estamos vendo que o conjunto de palavras tradicionalmente classificadas como "substantivos" e "adjetivos" não pode ser colocado em apenas duas classes sem deixar de exprimir pro-

priedades importantes dos itens envolvidos. Ou seja, dentro desse grupo de palavras, não se distinguem dois tipos de comportamento gramatical, mas pelo menos cinco (na verdade, como veremos, pelo menos sete).

Como se vê, não é possível dar uma resposta simples à pergunta "Como se devem definir sintaticamente os substantivos e os adjetivos?" Somos forçados a responder que o conjunto de itens que a gramática tradicional coloca nessas duas classes precisa ser colocado em um número maior de classes.

O trabalho de classificação dos itens que nos interessam não está realizado; no momento, há pelo menos um projeto em andamento nesse sentido, mas os resultados só serão conhecidos dentro de algum tempo. Assim, só se pode oferecer aqui uma solução provisória — suficiente, espero, para dar uma idéia da solução almejada. Essa solução provisória deverá ser substituída tão logo estejam disponíveis os resultados de uma pesquisa feita com maior número de traços e uma quantidade que represente uma amostra significativa do total dos "substantivos", "adjetivos" e também dos "pronomes", "numerais" e "artigos" da gramática tradicional.

Para dividir esse conjunto de itens em classes, dispomos até o momento de três traços (de segunda ordem), a saber, [NSN], [Mod] e [PN]. Outros traços podem ser propostos, descrevendo outros aspectos do comportamento gramatical das palavras. Por exemplo, podemos dividir as palavras entre as que podem ser "intensificadas" e as que não podem. Uma palavra é intensificada quando pode formar constituinte com um elemento em função de intensificador, como, por exemplo,

(18) Uma blusa [muito verde]

Em (18), *muito verde* é um constituinte (um sintagma adjetivo); e *verde* partilha esse sintagma com o intensificador *muito.* Esse traço pode ser notado assim: [Int—].

Outro traço que parece valer para os itens marcados [+Mod], e só para eles, é a propriedade de ocorrer como complemento do predicado (CP). O CP é muitas vezes preenchido por um SN, mas nem sempre; na frase (19), o CP não é um SN (*irritante* não poderia ocorrer sozinho como sujeito, por exemplo):

(19) Ernesto é irritante.

Faz sentido, portanto, definir um traço para as palavras que podem constituir, sozinhas, um complemento do predicado: [+CP].

Passemos a outro traço, que também pode ter relevância direta na delimitação das classes. A gramática tradicional aponta (corretamente, a meu ver) um grau de parentesco estreito entre os "substantivos" e alguns "pronomes", como *eu, você, ele, nós* — chamados, por isso, "pronomes substantivos". Esses itens podem ocorrer como núcleos de um SN, o que lhes

confere o traço [+NSN]; mas, além disso, não podem coocorrer com nenhum outro elemento do SN. Ou seja, formam, necessariamente, um SN por si sós:

(20) * O ele/* este ele/* nosso ele etc.

Essa é uma propriedade sintática importante e será considerada aqui um traço distintivo. Diremos que itens como *ele* são marcados [–T,SN], ou seja, **não têm** a propriedade de coocorrer com outro termo dentro do SN. Itens como *Patrícia* ou *inimigo*, naturalmente, serão marcados positivamente quanto a esse traço ([+T,SN]):

(21) Minha Patrícia/a Patrícia/aquela Patrícia

(22) Meu inimigo/o inimigo/aquele inimigo

Outro traço, finalmente, pode ser definido pela propriedade de desempenhar a função de predicativo; esse traço será abreviado [Pv].

Temos, portanto, sete traços sintáticos distintivos. A eles acrescentarei um traço de natureza morfológica (mas também formal, o que nos permite utilizá-lo ao lado dos demais), a saber, a propriedade de coocorrer com o sufixo de superlativo **-íssimo** (com suas variantes **-rimo** e **-imo**, como em *paupérrimo* e *facílimo*). Esse traço será representado pela notação [íssimo]. Chegamos a um total de oito traços distintivos e vamos basear neles nossa classificação do conjunto de itens habitualmente chamados "substantivos" ou "adjetivos" pela gramática tradicional.

Os oito traços nos permitem distinguir certo número de grupos de palavras, cada um deles definido por sua matriz característica; esses grupos serão a base para o estabelecimento das classes. Assim, o grupo a que pertencem as palavras *Patrícia, mesa* e *trabalho* é definido pela matriz [+NSN, +T,SN, –Pv, +CP, –Mod, –Int—, –PN, –íssimo]. Algumas das matrizes possíveis não correspondem a nenhuma palavra; por isso o total de grupos é bem menor do que o que se esperaria se se realizassem todas as combinações de traços.

Os grupos de palavras do tipo "substantivo ou adjetivo" que podemos vislumbrar como existentes na língua são sete, a saber:

(a) [+CP, –T,SN, –Pv, –Mod, +NSN, –Int—, –PN, –íssimo]
Exemplo: *ele.*

(b) [+CP, +T,SN, –Pv, –Mod, +NSN, –Int—, –PN, –íssimo]
Exemplo: *Patrícia.*

(c) [+CP, +T,SN, +Pv, +Mod, +NSN, +Int—, –PN, –íssimo]
Exemplo: *trabalhador.*

(d) [+CP, +T,SN, +Pv, +Mod, +NSN, +Int—, –PN, +íssimo]
Exemplo: *inimigo.*

(e) [+CP, +T,SN, +Pv, +Mod, +NSN, +Int—, +PN, +íssimo]
Exemplo: *verde.*

(f) [+CP, +T,SN, +Pv, +Mod, –NSN, +Int—, +PN, +íssimo]
Exemplo: *exato.*

(g) [+CP, +T,SN, +Pv, +Mod, –NSN, –Int—, –PN, –íssimo]
Exemplo: *presidencial.*

Esses grupos cobrem a quase totalidade dos itens tradicionalmente chamados "substantivos", "adjetivos" e "pronomes substantivos". Resta um pequeno grupo de "adjetivos" tradicionais (*mero, pretenso, simples* e alguns outros) que não cabem aqui; pertencem a pequenas classes fechadas.

Olhando a lista de sete grupos dada acima, não se percebe nada de parecido com a classificação tradicional em "substantivos", "adjetivos" e "pronomes". Mas parte da nomenclatura tradicional poderá ser útil para denominar classes de palavras derivadas dos grupos propostos; vamos a seguir discutir brevemente algumas conclusões que se pode tirar nesse sentido.

Uma observação interessante é a de que os traços [Pv] e [Mod] coincidem totalmente no corte que determinam no conjunto dos itens; isto é, todos os itens marcados [+Pv] são igualmente [+Mod], e todos os itens [–Pv] são também [–Mod]. Isso significa que há sempre uma diferença mínima de dois traços entre qualquer item dos grupos (a) ou (b) e qualquer item de qualquer dos outros grupos; ou seja, os itens parecem distribuir-se em dois grandes grupos, cuja distinção se faz através dos traços [Mod] e [Pv]. Cada um desses grandes grupos constituirá uma **classe**.

A primeira dessas classes compreende todos os itens marcados [+CP, –Mod, –Pv], isto é, os grupos (a) e (b) acima. E a segunda compreende os itens marcados [+CP, +Mod, +Pv], isto é, os grupos (c), (d), (e), (f) e (g). Como não há itens marcados [–Mod, +Pv], nem [+Mod, –Pv], essas duas classes são suficientes para englobar todas as palavras estudadas, sem resíduo. À primeira classe (definida como [–Mod, –Pv]), chamaremos **substantivos**; à segunda (definida como[+Mod, +Pv]), chamaremos **adjetivos**.

Substantivos são, pois, todas as palavras que podem ser complementos do predicado e não podem ser nem modificadores nem predicativos; adjetivos são as palavras que podem ser complementos do predicado, e também modificadores e predicativos. A inclusão nas definições do traço [+CP] é importante porque não queremos classificar como substantivo uma palavra como *apertamos.* Essa palavra (um verbo) não pode ser modificador nem predicativo; o que impede que seja classificada como substantivo é que ela não pode tampouco ser complemento do predicado. Isso vale para muitas outras palavras, como *ou, sempre, de, infelizmente* etc.: todas elas

são [–Mod, –Pv], mas não são substantivos porque são [–CP].

Os substantivos, por sua vez, se subdividem em duas subclasses, caracterizadas pelo traço [T,SN]. Uma das subclasses [o grupo (a)] se define como [+CP, –T,SN, –Pv, –Mod, +NSN, –Int—, –PN, –íssimo]; a outra subclasse [grupo (b)] se define como [+CP, +T,SN, –Pv, –Mod, +NSN, –Int—, –PN, –íssimo]. Podemos denominá-las, respectivamente, **substantivos tipo 1** e **substantivos tipo 2**. Os substantivos tipo 1 correspondem aproximadamente aos "pronomes substantivos" da gramática tradicional: *ele, aquilo, tudo.*

Mas essa correspondência é apenas aproximada; a subclasse dos substantivos tipo 2 engloba não só muitos "substantivos" tradicionais (*Patrícia, trabalho, mesa* etc.), mas ainda alguns "pronomes substantivos", como *alguém.* O leitor poderá verificar que *alguém,* ao contrário de *ele* e *aquilo,* pode ocorrer no SN sem ser sozinho, o que lhe confere o traço [+T,SN], e o coloca na classe dos substantivos tipo 2.

Seguindo esse tipo de raciocínio e procurando agrupar em subclasses grupos cujas matrizes sejam minimamente distintas, chega-se ao quadro abaixo, que aqui proponho como a classificação preliminar desse conjunto de itens:

Classes e subclasses de substantivos e adjetivos

Substantivos: [+CP, –Mod, –Pv]

Substantivos 1: [–T,SN, +NSN, –Int—, –PN, –íssimo]

Substantivos 2: [+T,SN, +NSN, –Int—, –PN, –íssimo]

Adjetivos: [+CP, +Mod, +Pv]

Adjetivos 1: [+Int—]

Adj. 1a: [–T,SN, +NSN, –PN, –íssimo]

Adj. 1b: [–T,SN, +NSN, –PN, +íssimo]

Adj. 1c: [–T,SN, +NSN, +PN, +íssimo]

Adj. 1d: [–T,SN, –NSN, +PN, +íssimo]

Adjetivos 2: [–Int—, –T,SN, –NSN, –PN, –íssimo]

Deve-se entender que cada subclasse possui todos os traços de sua classe; assim, os adjetivos 1a são também marcados [+Int—], porque pertencem à classe dos adjetivos 1, e [+CP, +Mod, +Pv], porque pertencem à classe maior dos adjetivos.

Em resumo, os substantivos compreendem duas subclasses (1 e 2); os adjetivos igualmente compreendem duas subclasses (1 e 2). Além disso, os adjetivos 1 ainda se subdividem em quatro subclasses (1a, 1b, 1c e 1d). Como se vê, a classificação proposta é mais complexa do que a da gramática tradicional; em compensação, descreve de maneira muito mais adequada o modo pelo qual esses itens funcionam dentro da língua.

A classe dos substantivos, como vimos, inclui alguns itens que a gramática tradicional não chama "substantivos". Da mesma forma, provavelmente a classe dos adjetivos inclui palavras que não se chamariam tradicionalmente "adjetivos", como *médico, homem, trabalhador.* Isso decorre, é claro, da maneira como definimos essas classes.

Não estou argumentando que a divisão em classes apresentada acima é a melhor possível; mas é fundamental que, uma vez definidas as classes, coloquemos cada item em sua classe seguindo estritamente a definição. Ora, *homem* pode ocorrer como modificador ou predicativo:

(23) Colimério tem um filho homem. (Mod)

(24) Eu o considero homem. (Pv)

Por conseguinte, precisa ser marcado [+Mod, +Pv], o que faz dele um adjetivo, segundo nosso sistema.

Os substantivos 1 são todos chamados "pronomes" pela gramática tradicional e poderão receber esse mesmo nome aqui; mas é preciso ter em mente que os **pronomes**, assim definidos, são uma subclasse dos substantivos; e que palavras como *alguém,* assim como todos os "pronomes adjetivos" da gramática tradicional, não são pronomes na nossa nomenclatura. A classificação dos "pronomes" tradicionais será considerada na próxima seção.

13.3. OS "PRONOMES" DA GRAMÁTICA TRADICIONAL

13.3.1. Exame da classificação tradicional

O grupo de itens que a gramática tradicional denomina "pronomes" não mostra traços comuns, nem sintáticos nem semânticos, que nos autorizem a colocá-los em uma classe única. Conseqüentemente, a classe tradicional dos "pronomes" terá de ser abandonada e substituída por diversas categorias; é o que passamos a ver.

A definição de "pronome" dada nas gramáticas, mesmo quando aplicável, nunca consegue delimitar exatamente o grupo de itens pretendido. Tomemos uma dessas definições, como exemplo típico:

> **Pronome** é a palavra que substitui ou acompanha o substantivo, indicando sua posição em relação às pessoas do discurso ou mesmo situando-o no espaço e no tempo. [Nicola & Infante, 1990, p. 201]

Vamos examinar essa definição detalhadamente.

Em primeiro lugar, há na definição dois termos de interpretação obscura: **substitui** e **acompanha**. A noção de "substituição", recurso freqüentemente usado para definir o "pronome", não é clara. Por exemplo, na frase seguinte

(25) Ela não gosta de quiabo.

o item *ela* pode ser um "substituto" para o substantivo *Giselle*, em certas circunstâncias. Mas igualmente, e nas mesmas circunstâncias, o substantivo *Giselle* pode ser substituído por outros elementos, que nunca são considerados "pronomes":

(26) Essa senhora não gosta de quiabo.

(27) Gigi não gosta de quiabo.

Por que não chamaríamos *senhora* ou *Gigi* de "pronomes"? Uma resposta que já encontrei foi a de que *Gigi* é um substantivo, logo não poderia ser pronome; mas essa resposta é circular, porque o que faz de um substantivo um substantivo, e não um pronome, é justamente o fato, alegado, de que não "substitui" coisa nenhuma. No mais, "substantivos" e "pronomes" (mais exatamente, "pronomes substantivos") teriam o mesmo comportamento sintático; isso é às vezes explicitado:

> Os **pronomes** desempenham na oração as funções equivalentes às exercidas pelos elementos nominais. [Cunha & Cintra, 1985, p. 268]

A noção de "acompanhamento" (o pronome "acompanha o substantivo") é, se não vaga, pelo menos inadequada. Um "pronome", *meu*, estaria acompanhando o substantivo em

(28) Meu velho paletó cinza

Mas é evidente que, nesse sentido, *velho* e *cinza* estão igualmente acompanhando o substantivo *paletó*, e no entanto não são "pronomes".

Temos de concluir, pois, que as propriedades de "substituir" ou "acompanhar" o substantivo não nos ajudam a delimitar a classe tradicional dos "pronomes". Ou seja, essa classe não apresenta coerência sintática visível.

O restante da definição vista é de natureza semântica. Mas ainda aqui, como veremos, a delimitação é seriamente inadequada. Segundo a definição, ao que parece, os "pronomes" indicariam a posição de um ser em relação às pessoas do discurso, ou então o situariam no espaço e no tempo.

É verdade que as palavras que situam seres em relação às pessoas do discurso (*nós*, *nosso* e talvez *este*, *esse*, *aquele*) são sempre chamadas "pronomes". Mas a maioria dos "pronomes" não têm nada a ver com as pessoas do discurso; assim, teriam de cair no segundo caso, situação no espaço e no tempo. Mas muitos não situam coisa alguma no espaço ou no tempo: *algum*, *qualquer*, *vários*, *que* etc. Além disso, há palavras que claramente situam seres no espaço ou no tempo, e nunca são chamadas "pronomes": *atual*, *antigo*, *contemporâneo* (situação no tempo); *próximo*, *distante*, *vizinho* (situação no espaço).

Essas considerações devem ser suficientes para convencer o leitor de que a definição examinada tem inadequações fundamentais e preci-

sa ser abandonada. Não temos tempo aqui para examinar outras definições, encontradas em outras gramáticas, mas é minha convicção que todas elas apresentam problemas comparáveis.

Impõe-se, na minha opinião, uma abordagem radicalmente diferente; a seguir, apresento uma nova classificação dos itens tradicionalmente chamados "pronomes", nos termos da análise proposta neste livro. Como veremos, a classe tradicional dos "pronomes" terá de desaparecer, pois se compõe na verdade de vários grupos, de comportamento sintático muito diverso.

13.3.2. Uma nova classificação

Deixando de lado a questão da classificação semântica (em si um problema de muito interesse), vou desenvolver agora uma classificação sintática dos itens tradicionalmente chama- dos "pronomes". Utilizarei apenas, com uma única exceção, traços já vistos nos capítulos anteriores.

(a) Em primeiro lugar, há um grupo de itens que se encaixam na classe dos **substantivos tipo 1**, estudados na seção 13.2. Como vimos ali, os substantivos tipo 1 se definem pelos traços [+CP, –Mod, –Pv, –T,SN, +NSN, –Int—, –PN, –íssimo]; ou, para abreviar, são itens que só podem ocorrer como complementos do predicado ou núcleos de um SN: [+CP, + NSN].

Os itens que compõem essa subclasse dos substantivos são todos chamados "pronomes" na gramática tradicional: *eu, ele, nós, eles, isto* etc. Propus na seção 13.2. que a designação **pronomes** fosse restrita a esses itens.

(b) Depois, há "pronomes" que devem ser classificados como **substantivos tipo 2**, ao lado de muitos itens tradicionalmente chamados "substantivos". Trata-se dos itens *alguém, algo, tudo,* que, como foi apontado em 13.2., diferem dos itens precedentes por poderem ocorrer no SN acompanhados por outro termo, como em *esse alguém, um alguém, alguém carinhoso; tudo de bom; algo misterioso.*

(c) Um terceiro grupo é o dos relativos e interrogativos: *que, quem* e *o qual* (este último, para efeitos de análise sintática, pode ser considerado um único item). Esses itens têm um comportamento todo especial, tanto no que se refere a sua posição na oração quanto no que se refere a suas propriedades semânticas; ver a respeito a seção 5.3.3.1. Sem entrar aqui em pormenores, vamos colocá-los em uma classe, caracterizando-os através do traço [+Rel] ("relativos"). Esse traço deverá ser explicitado como exprimindo suas propriedades sintáticas peculiares.

As três classes acima englobam os itens chamados tradicionalmente "pronomes substantivos". O que caracteriza seu comportamento, em nossos termos, é poderem todos eles ser núcleos de sintagmas nominais. Assim, pode-se afirmar que a distin-

ção tradicional entre "pronomes substantivos" e "pronomes adjetivos" tem alguma base sintática (naturalmente, a distinção é prejudicada pela ausência de base sintática que caracterize os "pronomes" como um grupo).

(d) Temos agora os itens *todos* e *ambos*, que se distinguem por serem os únicos que podem ser predeterminantes. Estes comporão uma classe por si sós, marcada [+PDet].

Agora temos os itens que ocupam a **área esquerda** do SN, estudada na seção 4.2.2. Conforme vimos ali, trata-se de um grupo de itens muito pequeno, mas de comportamento sintático complexo. Por isso, devem ser divididos em pelo menos quatro classes, a saber:

(e) Os que podem ser determinantes [+Det]: *o, um, algum, esse* etc.

(f) Os que podem ser possessivos [+Poss]: *meu, seu* etc.

(g) Os que podem ser quantificadores [+Qf]: *muitos, vários, único,* os ordinais etc.

(h) Os que podem ser numeradores [+Num]: *outro* e os cardinais.

Essas oito classes compreendem todos os "pronomes" da gramática tradicional. Naturalmente, os traços definitórios propostos não esgotam as propriedades sintáticas dos itens, e é possível subdividir várias das classes acima. Por outro lado, também é possível agrupar algumas em classes maiores, definidas por alguma propriedade mais geral; por exemplo, os itens das classes (a), (b) e (c), como vimos, têm em comum a propriedade de ocorrerem **sempre** como NSN e nisso se opõem aos das outras classes. Entretanto, não vou desenvolver a classificação além do ponto a que chegamos, pois isso demandaria mais tempo do que aquele de que dispõe um simples autor de gramáticas.

Uma observação importante é a seguinte: o grupo (b) acima representa uma classe aberta e muito extensa, onde os "pronomes" são uma pequena minoria; trata-se da classe dos substantivos tipo 2, que compreende a maioria dos "substantivos" tradicionais: dezenas de milhares de itens. Já os outros grupos representam classes fechadas, bastante pequenas. Os grupos (d), (f), (g) e (h) são compostos inteiramente de itens tradicionalmente chamados "pronomes"; e o grupo (d), definido como [+PDet], inclui não apenas "pronomes" (*aquele, este, alguns*), mas também os "artigos".

Quanto à nomenclatura das classes, duas delas já foram batizadas (**substantivos tipo 1** e **substantivos tipo 2**). Os itens do grupo (c) podem ser chamados **relativos** — tendo-se em mente que essa classe compreende igualmente os "interrogativos" da gramática tradicional. Quanto às outras cinco classes, prefiro não criar novos nomes para evitar uma proliferação excessiva de termos. Como cada uma dessas classes se define por um único traço,

podemos referir-nos a elas, sem ambigüidade, utilizando a própria definição. Proponho, pois, que se diga "**a classe [+Det]**", o que deve ser suficientemente claro, embora fira um pouco o estilo tradicional.

O quadro abaixo resume a classificação que proponho para os itens tradicionalmente chamados "pronomes":

> **Classificação dos itens tradicionalmente chamados "pronomes"**
>
> **Substantivos 1**: *eu, nós, ele, isto.*
>
> **Substantivos 2**: *alguém, tudo, algo.*
>
> **Relativos**: *que, o qual, quem.*
>
> **[+PDet]**: *todos, ambos.*
>
> **[+Det]**: *o, um, este, esse, aquele, alguns.*
>
> **[+Poss]**: *meu, seu, nosso.*
>
> **[+Qf]**: *muitos, vários, único, terceiro.*
>
> **[+Num]**: *outro, dois.*

13.4. PREPOSIÇÕES, CONJUNÇÕES, RELATIVOS E COORDENADORES

Vamos agora examinar um grupo de palavras de comportamento gramatical muito peculiar: essas palavras funcionam como elementos de conexão entre constituintes e são por isso chamadas **conectivos.** Trata-se das "preposições" e "conjunções" da gramática tradicional; a essas se acrescentam os relativos, já brevemente examinados na seção 13.3.

Distinguiremos dois tipos principais de conectivo; o primeiro tem como função sintática alterar a classe de um SN ou de uma oração — ou, mais precisamente, acrescentar-se a um SN ou a uma oração, formando um sintagma maior que pertence a outra classe que não SN ou O. A esses chamaremos **conectivos subordinativos**. O segundo tipo tem como função sintática juntar dois (ou mais) constituintes de mesma classe, formando o conjunto um constituinte maior que pertence à mesma classe dos constituintes conectados. Esses são os **conectivos coordenativos**.

A classificação aqui proposta serve para descrever o comportamento gramatical de grande parte dos conectivos. Restam muitos problemas, que ainda não estão devidamente pesquisados; darei uma rápida idéia desses problemas no final da seção.

13.4.1. Conectivos subordinativos

Vamos partir de um exemplo, o da palavra *de.* Sabemos que um SN não pode ser modificador; a função de modificador (interno ou externo) é típica de sintagmas adjetivos. A palavra *de* tem a propriedade de con

verter um SN em SAdj, dando-lhe, assim, a possibilidade de funcionar como modificador. É o que acontece no sintagma

(29) A mãe de Míriam

O item *Míriam* pode ser NSN e pode constituir, sozinho, um SN; mas não pode ser modificador. Já o sintagma *de Míriam* pode ser modificador; é, na verdade, um sintagma adjetivo. A função da palavra *de* é justamente a de formar, juntamente com um SN, um SAdj.

Em outros casos, pode-se converter um SN em sintagma adverbial. Por exemplo, um SN não pode desempenhar a função de adjunto circunstancial; mas um constituinte formado da palavra *em* mais um SN funciona como sintagma adverbial e pode ser AC:

(30) Míriam mora em Fortaleza.

Como se vê, palavras como *de* e *em* de certa forma "mudam" a classe e, portanto, as propriedades sintáticas dos SNs: *de* + *SN* é um SAdj, *em* + *SN* é um SAdv.

Há um grupo de palavras semelhantes a *de* e *em* que só podem construir-se com SNs, formando SAdjs e SAdvs; podemos chamá-las **preposições**. Definem-se assim:

Preposição *é a palavra que precede um SN, formando o conjunto um SAdj ou um SAdv.*

Exemplos são: *a, até, com, de, em, para, por, sem* etc.

Um segundo grupo de conectivos subordinativos é constituído de palavras que se acrescentam a orações, igualmente formando constituintes maiores de classe distinta: a seqüência forma um SAdv ou um SN. Por exemplo, seja a palavra *quando*:

(31) Míriam saiu quando começou a chover.

O constituinte formado de *quando* mais a oração *começou a chover* é um sintagma adverbial, cuja função é de adjunto circunstancial.

Outro exemplo é

(32) Míriam disse que carne de coelho faz mal.

Aqui a oração *carne de coelho faz mal* forma constituinte com *que* e o conjunto é um sintagma nominal, tanto é assim que funciona como objeto direto da oração (43).

Às palavras de comportamento semelhante ao de *quando* e *que* chamaremos **conjunções**; a definição é

Conjunção *é a palavra que precede uma oração, formando o conjunto um SAdv ou um SN.*

Em terceiro lugar temos palavras que se acrescentam a uma oração para formar um sintagma adjetivo. Aqui encontramos uma série de complicações notáveis. A língua não se contenta em fazer preceder uma oração de um conectivo para formar um SAdj; em vez disso, o próprio conectivo assume uma função sintá-

tica dentro da oração, que, correspondentemente, nunca tem um termo independente naquela função. Por exemplo,

(33) Míriam vai comprar a casa que você recomendou.

A oração subordinada aqui é *que você recomendou.* E vimos, em 5.3.3.1., que o elemento *que* deve ser analisado como objeto direto dessa subordinada; ele de certo modo tem a função extra de "representar" na subordinada o substantivo *casa*, que é o núcleo do objeto direto da oração principal. A esses conectivos chamamos **relativos**. Os relativos são: *que, o qual, quem, cujo, onde*, além de *quanto*, que é bastante raro (só ocorre na construção *tudo quanto você quiser*).

Preposições, conjunções e relativos são as três classes de conectivos subordinativos existentes na língua.

13.4.2. Conectivos coordenativos

Vejamos agora o funcionamento da palavra *e*. Se ligarmos dois SNs por meio de *e*, teremos um conjunto que funciona, ele próprio, como um SN. Assim, temos:

(34) Míriam e Thiago fazem Belas-Artes.

(35) Eu conheço bem Míriam e Thiago.

A seqüência *Míriam e Thiago* funciona como um SN: sujeito em (34) e objeto direto em (35).

Igualmente podemos ligar dois SAdjs, dois SAdvs etc., formando respectivamente um novo SAdj e um novo SAdv.:

(36) Casas de muros brancos e de telhado vermelho

(37) Já morei em Fortaleza e em Natal.

Além disso, *e* também pode ligar constituintes internos dos sintagmas; por exemplo, em

(38) O meu e nosso amigo Marcão

Aqui *e* está ligando *meu* e *nosso*, formando um constituinte que é, ele próprio, o possessivo do sintagma (38).

Às palavras que funcionam como *e* chamaremos **coordenadores**; definem-se da seguinte maneira:

Coordenador *é uma palavra que liga dois constituintes de mesma classe, formando o conjunto um constituinte da mesma classe que os dois primeiros.*

Os coordenadores mais típicos (e também os mais bem comportados) são *e* e *ou*. Outras palavras tradicionalmente classificadas como "conjunções coordenativas" se assemelham a essas, mas apresentam também diversas idiossincrasias que estão por estudar: *mas, pois, nem, que, porque* e várias outras.

13.4.3. Algumas notas sobre os conectivos

13.4.3.1. Complemento complexo de preposições

Vimos que as preposições se acrescentam a SNs; e vimos que certas conjunções se anexam a orações para formar SNs. Daí resulta que os dois processos podem ocorrer na mesma construção, resultando em uma preposição acompanhada de complemento complexo. Assim, *que* mais uma oração forma um SN, como em

(39) ... que estivesse chovendo.

Esse SN pode por sua vez receber uma preposição, formando o conjunto, digamos, um SAdv. É o que acontece na frase

(40) Ela usava guarda-chuva sem que estivesse chovendo.

Essa possibilidade é uma decorrência da análise de *que* + oração como formando um SN. Há restrições à montagem de tais sintagmas complexos com preposição; provavelmente muitas são restrições de ordem semântica — por exemplo, não encontramos * *para que estivesse chovendo*, porque há incompatibilidade entre *para* (finalidade) e *estar chovendo* (evento natural, não-dependente de intenção).

13.4.3.2. Preposições compostas

Há seqüências de duas (às vezes três) palavras que se comportam como preposições e se analisam tradicionalmente como tais: *antes de, apesar de, através de, junto a, para com, de acordo com* etc. Sem entrar a fundo na questão, vou aceitar essa análise, admitindo que uma preposição pode ser composta de duas ou três palavras. Assim, uma seqüência como *antes de* será um único item léxico, a saber, **uma preposição composta**.

13.4.3.3. Preposições e conjunções homônimas

As gramáticas costumam citar bom número de preposições e conjunções homônimas (e, além do mais, sinônimas), como: *antes de* (prep.)/ *antes que* (conj.); *depois de* (prep.)/ *depois que* (conj.); *desde* (prep.)/*desde que* (conj.) etc. Essa é uma situação suspeita: por que haveria tantos pares de preposição e conjunção tão semelhantes em forma e significado?

Temos agora uma maneira bem simples de explicar o fenômeno. Tomemos *desde*; em (41), é claramente uma preposição:

(41) Meu pai trabalha desde a infância.

Já em (42) a gramática tradicional consideraria *desde que* uma conjunção:

(42) Meu pai trabalha desde que era criança.

No entanto, sabemos que *que* + oração formam um SN; e um SN pode construir-se com uma preposição, para formar um sintagma adverbial. É exatamente o que acontece em (42): a preposição (não conjunção) *desde* se coloca antes do SN (não oração) *que era criança*, formando um sintagma adverbial que funciona, na oração (42), de maneira paralela à do SAdv *desde a infância* em (41).

A única complicação que surge é que, quando a preposição é composta com o elemento *de* como segunda parte (*antes de, depois de*), esse *de* em certos casos se omite antes de conjunção. Assim, temos

(43) Meu pai entendeu tudo antes da explicação.

(44) Meu pai entendeu tudo antes que você explicasse.

Não se pode dizer * *antes de que você explicasse*. Essa omissão de *de* não se limita a casos de preposição composta; quando *de* ocorre independentemente, acontece o mesmo:

(45) Meu pai gostou de você.

(46) Meu pai gostou que você viesse.

e não * *de que você viesse*.

A supressão de preposição antes de conjunção não é sempre obrigatória; com certas preposições compostas, o *de* se mantém. A preposição pode ser mantida quando introduz um modificador, ainda que este comece por conjunção (sendo, portanto, oracional):

(47) a. Tenho medo de que apareça alguém.

No entanto, a omissão da preposição me parece aceitável mesmo nesses casos:

b. Tenho medo que apareça alguém.

13.4.3.4. Conectivos e regência

Uma propriedade dos conectivos subordinativos é poderem determinar o modo da oração com que se constroem. Assim, podemos contrastar

(48) O governo renunciou depois que se revelou o escândalo.

(49) O governo renunciou antes que se revelasse o escândalo.

A preposição *antes* exige subjuntivo na oração que lhe é subordinada, ao passo que *depois* admite o indicativo ou o subjuntivo; *depois* ocorre com subjuntivo em

(50) O governo só renunciaria depois que se revelasse o escândalo.

Como se vê, parece que com *depois* o modo da subordinada depende de regras de correlação modo-temporal entre a subordinada e a principal; ou seja, a preposição propriamente dita não determina o modo da subordinada. Já com *antes* nunca se pode usar o indicativo.

Este é mais um dos muitos pontos que ainda não receberam um estudo cuidadoso; por isso, ficará aqui apenas esta rápida nota.

13.4.3.5. Preposições com sintagma adverbial

Finalmente, é preciso observar que em certos casos é possível acrescentar uma preposição a um sintagma adverbial, mesmo se este se inicia por outra preposição. Assim, podemos dizer

(51) Ela sempre passa por aqui.

(52) O avião passou por sobre a cidade.

(53) Ela saiu de dentro do tonel.

(54) Ela saiu dali.

Nem todas as combinações são admissíveis: não se aceita * *com em*, * *sobre por* nem * *diante de contra.* Falta fazer um levantamento completo, assim como interpretar devidamente as restrições: até onde vai a ação das incompatibilidades semânticas, e até onde a das marcas sintáticas?

13.5. ADVÉRBIOS

13.5.1. Existe uma classe dos "advérbios"?

A categoria tradicional dos "advérbios", assim como a dos "pronomes", encobre uma série de classes, às vezes de comportamento sintático radicalmente diferente. Ao contrário dos "pronomes", porém, os "advérbios" do português estão muito pouco estudados em seu conjunto; temos apenas estudos parciais. A situação é tal que não parece possível dar uma visão abrangente das diversas classes, nem mesmo uma lista completa delas. Aqui vou limitar-me a alguma exemplificação, pois a pesquisa que permitirá uma descrição mais detalhada está por fazer.

Naturalmente, temos que definir cada classe em termos de seu potencial funcional. A definição tradicional fala da propriedade de "modificar" itens de outras classes — ou mesmo de "modificar o próprio advérbio", o que introduz na definição um elemento de circularidade que a inviabiliza. Embora não seja esse o nosso critério, terei algo a dizer a esse respeito mais adiante.

Como esta seção é totalmente programática, oferecendo não uma análise, mas sugestões de análise, vamos tomar apenas um pequeno grupo de palavras tradicionalmente chamadas "advérbios", para ilustrar suas

diferenças de potencial funcional e, portanto, de classe. Sejam os itens

não, rapidamente, completamente, muito, francamente.

Essas cinco palavras são todas classificadas como "advérbios" e subclassificadas segundo um critério semântico ("de negação", "de modo", "de intensidade") que não podemos levar em conta dentro de um estudo sintático. É óbvio que a classificação semântica tem também interesse, mas precisa ser realizada, segundo os princípios aqui aceitos, separadamente. Aqui nos ocuparemos da classificação morfossintática desses cinco itens.

As cinco palavras da lista podem ser encontradas desempenhando diversas funções sintáticas, a saber:

Negação verbal:

(55) Seu tio *não* apareceu na estação.

Intensificador:

(56) Almeida é *muito* magro.

(57) Almeida estava *completamente* bêbado.

(58) Essa proposta é *francamente* ilegal.

Adjunto circunstancial:

(59) Ela ri *muito.*

Atributo:

(60) Terminamos a pintura *rapidamente.*

(61) Ela me revelou tudo *francamente.*

Adjunto adverbial:

(62) Ela decorou o apartamento *completamente.*

Adjunto oracional:

(63) *Francamente,* acho que ele nos enganou.

Como se vê, essas palavras podem ocupar pelo menos seis funções; e algumas podem ocupar mais de uma função. Dos exemplos acima, podemos tirar um primeiro esboço do potencial funcional de cada uma (as funções não mencionadas são marcadas negativamente):

não [+NV]
rapidamente [+Atr]
completamente [+Int, +AA]
muito [+Int, +AC]
francamente [+Int, +AO, +Atr]

Como se vê, obtemos cinco potenciais funcionais distintos. Poderíamos ver aí, por conseguinte, cinco classes dentro do grupo tradicional dos "advérbios". Já aí se pode observar como a classificação tradicional deixa de exprimir as diferenças encontradas entre esses itens, no que diz respeito a seu comportamento gramatical. Note-se, aliás, que a subclassificação em advérbios "de modo", "de intensidade" etc. não corresponde à classificação sintática obtida acima: há ali três "advérbios de modo" (*rapidamente, completamente* e *francamente*), que no entanto são sintaticamente bem diferentes.

Os traços que caracterizam as presumíveis cinco classes de advérbios não são todos exclusivos das palavras tradicionalmente assim classificadas. Assim, o traço [+AC] é partilhado por um grupo de adjetivos do tipo 1c (ver a seção 13.2.), como *alto, baixo, rápido, fundo* e vários outros. Esses adjetivos podem ser adjuntos circunstanciais, como em

(64) Ela escreve rápido.

(65) Joaquim mergulhou fundo na questão.

Têm, portanto, o traço [+AC], assim como *muito.* Por outro lado, diferem de *muito* por não poderem ser intensificadores e por poderem ser pré-núcleos e modificadores.

A classificação dada acima aos cinco itens sugere que não será fácil estabelecer uma classe que abarque a totalidade ou a maioria dos itens tradicionalmente chamados "advérbios". Pessoalmente, acho perda de tempo tentar justificar assim a análise tradicional, simplesmente porque, neste ponto, ela é claramente inadequada. Como, a bem dizer, a pesquisa detalhada ainda não se fez, isso não passa de um palpite — a ser devidamente validado ou invalidado, conforme for o caso.

Por ora, portanto, tendo a afirmar que não existe uma classe que compreenda, mesmo aproximadamente, os itens tradicionalmente chamados de "advérbios". As diferenças sintáticas entre os "advérbios" são muito profundas, em parte comuns a palavras de outras classes tradicionais, e não autorizam a postulação de uma classe única. Temos aqui, na verdade, diversas classes, que podem sem dúvida agrupar-se, mas dificilmente de maneira análoga à proposta pela análise tradicional.

13.5.2. "Modificação"

A definição tradicional diz que o advérbio "modifica" determinadas classes (entre as quais o próprio advérbio, como vimos). E muitas análises mais modernas retomam basicamente essa noção, de modo que vale a pena examiná-la aqui.

A noção de "modificação" é bastante obscura; como a interpreto, seria um misto de semântica e sintaxe. Semanticamente, "modificação" significa que um advérbio teria seu significado amalgamado ao de um outro elemento, formando um todo semanticamente integrado; assim, digamos, *corremos* exprime uma ação, e *corremos depressa* exprime a mesma ação, acrescida de algum ingrediente de significado. Tanto *corremos* quanto *corremos depressa* seriam unidades no plano semântico.

Essa observação, embora vaga, é provavelmente correta; mas não nos ajudará a caracterizar o "advérbio", porque se aplica a outras classes. Assim, *comi* é uma ação, e *comi uma peixada* é a mesma ação, acresci-

da de um ingrediente semântico que a especifica melhor.

Sintaticamente, a noção de "modificação" parece referir-se à ocorrência conjunta dentro de um constituinte: o que se chama em sintaxe **estar em construção com**. Desse modo, *corremos depressa* forma um constituinte (*corremos* está em construção com *depressa*). Isso, por si só, não é suficiente para definir o "advérbio", porque é claro que *comi* e *uma peixada* também estão em construção em *comi uma peixada*. Se definíssemos o advérbio como o elemento que ocorre em construção com um verbo, *uma peixada* teria de ser um constituinte adverbial.

Além disso, em certos casos, não é fácil dizer de maneira geral com o que determinado termo está em construção. Tomemos o caso de um atributo. Em certas posições, pode-se sustentar que o atributo está em construção com o NdP, como em

(66) Terminamos *rapidamente* a pintura.

Mas em outros casos isso é difícil de defender:

(60) Terminamos a pintura *rapidamente*.

(67) *Rapidamente* os operários terminaram a pintura.

Na verdade, parece que "estar em construção" com este ou aquele termo não é uma propriedade fundamental das funções sintáticas. Encontramos atributos (assim como várias outras funções) em posições tão variadas na oração que a solução mais prudente, adotada aqui, é considerá-lo simplesmente um constituinte de nível oracional: está em construção com todos os demais constituintes de nível oracional, para formar a oração.

A tentativa de definir o advérbio em termos do elemento que ele "modifica" é adotada por autores não-tradicionais, mas creio que eles o fazem por falta de melhor alternativa, na ausência de estudos abrangentes. Assim, Huddleston, em um trabalho em geral excelente, afirma que

> para os nomes, os modificadores mais característicos são adjetivos, enquanto que para os verbos são advérbios. [1984, p. 330]

Neste ponto, o autor se deixou levar pela noção tradicional, que é a meu ver insustentável. Primeiro, a afirmação é inverificável na prática: Huddleston afirma que com verbos o que ocorre são em geral os advérbios; mas a definição de "advérbio" depende da ocorrência como modificador de verbo. Em outras palavras, "ser advérbio" é, em parte, "ser modificador de verbo"; nesse contexto, a observação de que os modificadores de verbos são em geral advérbios carece de sentido independente.

Em segundo lugar (e admitindo uma concepção informal de "advérbio" como membro de uma lista de palavras tradicionalmente assim chamadas), há os casos presumivelmente menos característicos, que

também teriam de ser levados em conta. Por exemplo, em

(68) Somente André percebeu a situação.

temos uma palavra, *somente*, que é em geral considerada advérbio, modificando um substantivo, *André*. Por outro lado, em

(64) Ela escreve rápido.

rápido, que provavelmente deve ser considerado um adjetivo, está modificando um verbo.

Como se vê, há poucas esperanças de se chegar a uma definição adequada de qualquer classe em termos dos elementos que ela "modifica". A conclusão é que é necessário lançar mão das funções, tal como fizemos na seção precedente. A diferença entre *uma peixada* e *depressa* é que o primeiro elemento é objeto direto, e o segundo é adjunto circunstancial. Assim, "estar em construção com um verbo" (minha tradução de "modificar um verbo") não caracteriza um advérbio frente a um sintagma nominal; o que os diferencia claramente são as diferentes funções que desempenham quando estão em construção com o verbo.

Conclui-se, então, que o fato de estar em construção com o verbo, ou com o adjetivo etc. não pode ser utilizado como critério definitório de nenhuma classe. A definição de "advérbio", se for possível (o que duvido), deverá ser formulada em termos de funções. Por ora, ficaremos com a idéia de que sob o rótulo de "advérbio" se esconde uma variedade irredutível de classes.

Que classes são essas e como se definem é uma pergunta que ainda não foi, que eu saiba, satisfatoriamente respondida. Na verdade, a taxonomia dos chamados "advérbios" é uma das grandes áreas inexploradas da gramática portuguesa. Não têm faltado pesquisas sobre o tema, mas em geral não consideram os dados em sua totalidade, ou partem de pressupostos que considero inadequados. Falta, acima de tudo, uma síntese. Para uma primeira aproximação, o leitor interessado poderá ler Ilari et alii, 1990 — um dos poucos trabalhos onde se tenta uma abordagem ampla da questão.

14 O léxico

Vimos no capítulo 2 que o léxico é uma longa lista onde se armazena a informação idiossincrática (ou seja, não redutível a regras gerais) da língua. O léxico é composto de grande número de itens (**itens léxicos**), e cada um deles encerra informação sobre as características fonológicas, morfológicas, sintáticas e semânticas de uma palavra (melhor dizendo, de um **lexema** — ver adiante), de um morfema, ou ainda de uma expressão idiomática.

14.1. CARÁTER DO LÉXICO

Para compreender o papel do léxico dentro da gramática, consideremos o que significa saber uma língua. Para que alguém saiba uma língua, é necessário que domine (na prática) uma série de regras; assim, para saber português, é preciso saber que não se pronuncia um [o] (como em *avô*) no final de uma palavra se esta não for oxítona; ou seja, o que se ouve no final de *Pedro* é antes um [u]. É preciso também saber que se pode dizer *escrevi, escreveu, escrevendo, escrevemos,* mas não * *escrevinha,* *es*crevemente,* * *escrevíssimo.* Outra coisa que se deve aprender é que se pode dizer *O livro caiu nesse buraco,* mas não * *Livro o caiu buraco nesse.*

Todos esses conhecimentos são expressos por **regras**, ou seja, instruções gerais para a construção dos enunciados da língua portuguesa. Cada uma das três regras mencionadas vale para grande número de itens, desde que satisfaçam às condições expressas na própria regra; assim, a proibição de pronunciar [o] vale para qualquer vogal átona final de palavra; a proibição de acrescentar o sufixo *-mente* vale para qualquer forma de verbo; e a seqüência * *livro o* é proibida por regras que estabelecem que o artigo deve vir sempre antes do substantivo, nunca depois dele. Justamente por se aplicarem a muitos itens, sendo estes definidos de maneira geral, é que se chamam **regras**.

Além desse tipo de conhecimento, constituído por regras mais ou menos gerais, saber uma língua

também implica o conhecimento de grande número (vários milhares) de elementos individuais, cuja forma e significado não são governados por regras. É o caso, notadamente, das **palavras** da língua: além de conhecer as regras fonológicas, morfológicas, sintáticas e semânticas do português, precisamos conhecer a **lista das palavras** do português.

Nenhuma regra nos ajudará a saber o que significa, digamos, *enólogo*. É necessário conhecer a própria palavra (ou, o que é equivalente, os elementos constitutivos da palavra, pois esta se compõe de *eno-* mais *-logo*). E conhecer a palavra implica em conhecer sua pronúncia (começa com [e], depois vem um [n] etc.); suas propriedades morfossintáticas (pode ser núcleo de um sujeito, pode ocorrer depois de um artigo); e suas propriedades semânticas (designa uma pessoa que se ocupa do estudo ou da produção do vinho).

A diferença entre o conhecimento regular e o conhecimento particularizado fica bem clara quando consideramos as pseudopalavras *epólogo* e *tpelóg*. Do ponto de vista das regras do português, *epólogo* é perfeitamente bem formada: nenhuma regra proibiria a existência de tal palavra, ao contrário de *tpelóg*, que não é **nem poderia ser** uma palavra do português. Acontece que nem *epólogo* nem *tpelóg* existem em português; mas *tpelóg* não existe porque é fonologicamente mal formada, ao passo que *epólogo* não existe por acidente: nenhuma razão gramatical impediria alguém de introduzir essa palavra, caso viesse a ser necessária. Já *tpelóg* nunca seria aceita pelos falantes da língua, pois sua estrutura vai contra as regras fonológicas do português. Ninguém precisa consultar um dicionário para saber que *tpelóg* não é uma palavra portuguesa.

Como todos sabemos, o estudo de uma língua inclui sempre a memorização de grande número de palavras (e também de grande número de morfemas). A lista dessas palavras ou morfemas se denomina **léxico**, e o léxico pode ser considerado, *grosso modo*, como um repositório da informação idiossincrática (não diretamente governada por regras) da língua. O léxico não se compõe apenas de palavras, pois compreende também morfemas separados, além de expressões fixas do tipo *a olhos vistos, bater as botas* etc. Neste livro nos interessam em especial as palavras; mas ao se fazer o estudo da morfologia da língua é necessário considerar também os morfemas separadamente.

A informação que cada item léxico contém, como vimos, relaciona o item com todos os componentes da gramática: é fonológica (como se pronuncia); morfológica (composição em morfemas, possibilidades de variação); sintática (como se insere nas orações); e semântica (o que significa).

É por isso que não é correto colocar o estudo do léxico (como, por exemplo, sua divisão em classes

de palavras) na morfologia, como fazem as gramáticas tradicionais. O estudo do léxico é uma parte do estudo da língua que não se insere propriamente na gramática, embora esteja intimamente ligado a ela. Esse estudo se denomina **lexicologia**. Em outras palavras, o léxico não pertence a nenhum dos componentes da gramática (fonológico, morfológico, sintático, semântico); antes, é o lugar onde todos os componentes se cruzam, depositando a informação idiossincrática.

14.2. LEXEMAS, PALAVRAS, MORFEMAS E EXPRESSÕES IDIOMÁTICAS

Costuma-se usar o termo **palavra** para designar uma forma individual como *pedra* e também para designar um conjunto de formas relacionadas de determinada maneira, como o par *pedra/pedras*. Agora será necessário procurar um uso mais rigoroso dos termos; teremos de perguntar: *pedra* e *pedras* são duas palavras diferentes ou são duas formas da mesma palavra?

Pedra e *pedras* precisam ser reunidas em uma entidade única para efeitos de armazenamento no léxico; com efeito, a relação entre essas duas formas é sistemática, e não faria sentido armazená-las independentemente. Por outro lado, não há dúvida de que são formas diferentes e, para certos fins, é conveniente poder distingui-las. O que precisamos é, evidentemente, duas noções distintas, embora relacionadas.

Vamos, pois, denominar **palavra** cada uma das formas individuais *pedra* e *pedras*; ou seja, temos em português a palavra *pedra* e a palavra *pedras*. E à entidade que engloba *pedra* e *pedras* daremos o nome de **lexema**: as **palavras** *pedra* e *pedras* são duas formas do **lexema** *pedra*. Podemos dizer, então, que uma palavra pode ser singular ou plural; mas um lexema pode ter um singular e um plural. E diremos também que o léxico é (entre outras coisas) **a lista dos lexemas da língua**.

A relação entre as diversas palavras que formam um lexema é regular, sistemática; para ser mais específico, direi que a relação entre as palavras membros de um lexema é puramente **flexional**: elas são formas flexionadas do lexema. Assim, *pedra* é o singular, e *pedras* o plural do lexema *pedra*. Podemos definir o lexema assim:

> Lexema *é um conjunto de palavras que diferem apenas quanto a morfemas flexionais.*

Essa é a definição dada em Perini, 1985 (p. 27-36), para o "paradigma"; entretanto, esse uso do termo *paradigma* tem causado certa confusão, motivo pelo qual prefiro adotar agora o termo *lexema* (estou aqui seguindo o uso mais freqüente em lingüística).

A definição depende, naturalmente, de se poder distinguir claramente os morfemas flexionais dos não-flexionais (isto é, derivacionais); para uma discussão, remeto o leitor à passagem mencionada de meu livro de 1985. Aqui darei apenas uma rápida explicação.

Chama-se **flexão** uma variação da forma (e do significado) da palavra que se caracteriza por grande sistematicidade de ocorrência. Por exemplo, *pedras* é uma flexão de *pedra* porque praticamente toda palavra da classe de *pedra* se associa a uma forma de plural. O mesmo acontece com as formas de um verbo, como *pegar, pego, pegando, peguei* etc. — toda palavra da classe de *pegar* se associa a um conjunto formal e semanticamente paralelo a esse (todo verbo tem presente, infinitivo e pretéritos; tem primeira e terceira pessoas etc.). Dizemos, então, que em português existe o lexema *pegar* (ou, mais simplesmente, o verbo *pegar*), e que esse lexema apresenta diversas formas flexionadas, cada uma das quais é uma palavra: *pego, peguemos, pegando* etc.

A **derivação**, por outro lado, não possui essa sistematicidade. Podemos relacionar *fazer* e *desfazer*, mas nem todo verbo tem uma forma associada em *des-*: * *deschegar*, * *desdormir*, * *desescrever* etc. Podemos formar, derivada de *livro*, a palavra *livresco*, significando algo que se relaciona a livros de certa maneira. Mas nem toda palavra da classe de *livro* (ou mesmo do campo semântico de *livro*) pode produzir algo comparável: * *escolesco*, * *cinemesco*, * *teatresco*, * *folhetesco* etc. Assim, *livro* é um lexema, e *livresco* outro lexema separado; o mesmo se dá com *fazer* e *desfazer*.

A nomenclatura tradicional reconhece implicitamente essa distinção, pois comumente nos referimos a *pedras* como "uma forma da palavra *pedra*", mas não dizemos que *desfazer* é uma forma do verbo *fazer*, nem que *livresco* é uma forma de *livro*. Sem dizer que a distinção entre flexão e derivação seja isenta de problemas, terá de servir para basear nossa noção de "lexema".

O léxico, portanto, inclui uma lista de lexemas: *fazer, desfazer, pedra, eu, sempre, livro, livresco*, mas não, por exemplo, *faço* (que é uma flexão de *fazer* e é, portanto, especificado na gramática, através de regras, paradigmas, exceções etc.). E cada lexema constitui, no léxico, um item léxico.

Além disso, o léxico precisa listar formas que não são palavras e que igualmente constituem itens léxicos. Não nos ocupamos desses itens neste livro, pois aqui não se tratou de morfologia; mas é importante mencioná-los, para que não se tenha a impressão de que todo item léxico é necessariamente composto de um lexema, e este de uma ou mais palavras.

O falante do português tem a possibilidade (algo limitada) de formar lexemas a partir de constituintes que podem ser menores do que a palavra, chamados **morfe-**

mas. Assim, podemos usar o termo *canivetinho*, mesmo que nunca o tenhamos ouvido ou lido antes; a nova palavra será compreendida sem dificuldade por qualquer falante do português. Nós a compusemos a partir de *canivete* e do morfema *-inh(o)*, que nunca ocorre independentemente como uma palavra.

Esses "pedaços de palavra" se chamam **morfemas presos** (presos porque nunca ocorrem como palavras, que são formas livres). Evidentemente, eles fazem parte de nosso acervo lingüístico e precisam ser listados no léxico, como itens independentes. Aqui não entraremos em pormenores sobre como se inserem tais elementos no léxico; basta observar que também eles têm seu lugar.

Finalmente, o léxico precisa incluir ainda certas expressões idiomáticas fixas, do tipo *bater as botas, a olhos vistos* etc. Estas não são propriamente palavras: por exemplo, em *bater as botas*, podemos flexionar a primeira parte: *bateu as botas, baterão as botas* etc., o que nunca acontece com as palavras propriamente ditas. Mas as expressões idiomáticas não podem tampouco ser consideradas frases ou sintagmas normais, por várias razões: primeiro, na fala, nunca podem ser interrompidas por hesitações, sem destruir o efeito de expressão idiomática. Assim, se alguém disser

(1) Cidinha bateu — ééé... — as botas.

a mensagem transmitida não será de que "Cidinha morreu", mas de que realmente bateu umas botas (para tirar a poeira, talvez).

Depois, essas expressões se compõem de elementos fixos. Não podemos sequer mudar certas flexões: a frase

(2) Cidinha bateu a bota.

novamente quer dizer que ela esteve limpando o calçado, e não que morreu.

Finalmente, em certos casos, como em *a olhos vistos*, a própria estrutura da expressão é peculiar e não corresponde exatamente à estrutura de um sintagma (igualmente, se interpretarmos a expressão literalmente, o resultado pode ser anômalo).

Concluímos que tais expressões não são estruturas montadas pela sintaxe e interpretadas pela semântica, mas verdadeiros itens compostos, listados separadamente no léxico.

14.3. O ITEM LÉXICO

14.3.1. Matrizes de traços

Vimos que o item léxico armazena a informação idiossincrática acerca dos lexemas da língua. Essa informação pode ser de vários tipos, de maneira que o item léxico compreende diversas **matrizes** de traços, cada uma delas fornecendo informação relacionada a um dos componentes da gramática.

Tomemos o exemplo do lexema *casa*. A informação a ser colocada no léxico a respeito desse lexema deve especificar pelo menos o seguinte:

— sua pronúncia;

— suas possibilidades de variação (flexões e derivações);

— suas propriedades sintáticas;

— seu significado.

Respectivamente, o item léxico incluirá traços fonológicos, traços morfológicos, traços sintáticos e traços semânticos.

A matriz de traços fonológicos fornece instruções para a pronúncia de *casa*; essa matriz é abreviada através de uma **transcrição fonológica**:

/'kaza/.

(O acento inicial indica a posição da sílaba tônica.)

Nem todos os detalhes da pronúncia são indicados na transcrição fonológica, porque alguns são previsíveis e constituem, portanto, regras do componente fonológico. Assim, não é necessário indicar no léxico que os dois **aa** de *casa* se pronunciam diferentemente (o último é sensivelmente mais fechado do que o primeiro). Isso é conseqüência automática de uma regra que estabelece que um **a** átono final é sempre pronunciado mais fechado. Ou seja, essa pronúncia mais fechada do **a** átono final não é uma idiossincrasia do item *casa*, e portanto não consta de sua explicitação.

A matriz morfológica informa acerca das possibilidades de variação do item. Também aqui pode haver traços idiossincráticos e traços regulares. A possibilidade de formar o plural em *-s* não deve ser colocada no item, porque é característica regular de todos os substantivos (e *casa* é marcado como substantivo na matriz sintática); mas, se se tratasse do item *mal*, seria necessário estipular que seu plural é *males* e não **mais*, que seria a forma regular.

A matriz sintática fornece a classificação do item quanto a seus traços sintáticos. Para *casa*, temos traços como [+T,SN, +NSN, –Int—, –PN, +PS, –Mod, –Pv] (ver a significação desses traços no capítulo 13), o que caracteriza esse item como um **substantivo tipo 2**. Além disso, *casa* é feminino — ou seja, exige determinada forma de certos elementos que o acompanhem no mesmo SN: *a* (e não *o*); *amarela* (e não *amarelo*) etc.

Finalmente, a matriz semântica dá o significado básico do item. *Casa* é concreto, inanimado (e portanto não-humano); significa um local, esse local serve como habitação, e assim por diante.

Todos esses tipos de informação são essenciais para que se possa inserir o item *casa* corretamente nos enunciados do português. Se violarmos os traços da matriz fonológica (o que, obviamente, só se pode fazer na fala), poderemos produzir algo como

(3) * Mandei pintar minha cosa.

Não será possível reconhecer o item *casa* porque a pronúncia está incorreta.

Desobedecendo as instruções da matriz morfológica, diríamos

(4) * No meu bairro há mais de cem cásaes.

Fizemos o plural de *casa* acrescentando *-es* (como se faz, corretamente, com *mal*); novamente, o resultado é inaceitável.

Se desobedecermos a matriz sintática, poderemos produzir algo como

(5) * Aluguei um barracão casa.

onde *casa* está usado como modificador (apesar de ser marcado [–Mod]).

Finalmente, podemos violar os traços semânticos, dizendo coisas como

(6) * Minha casa está querendo um aumento de salário.

onde tratamos *casa* como se fosse humano.

Observe-se que algumas dessas frases podem ser utilizadas em contextos especializados (fábulas, poesia etc.). São inaceitáveis enquanto ocorrências na fala (ou escrita) comum; mas sabemos que violar a gramática é um dos recursos mais usados na linguagem poética.

14.3.2. Delimitação

O estudo do léxico é prejudicado por um fato bastante incômodo: não existe um critério seguro que nos permita saber, em certos casos, onde termina um item léxico e onde começa outro. Esse problema surge quando temos casos de homonímia ou de polissemia. Vimos na seção 9.4.2. que não se sabe bem como distinguir casos de polissemia de casos de homonímia: frente a uma forma como *folha*, como decidir se se trata de um item léxico com dois significados (polissemia) ou de dois itens léxicos que se pronunciam e escrevem da mesma maneira (homonímia)?

Como resultado, a estrutura do léxico fica até certo ponto indefinida. Se quisermos fazer uma estatística do número de itens que têm determinados traços — algo que é necessário fazer para estabelecer a importância relativa das classes de palavras —, teremos de contentar-nos sempre com números aproximados. A questão da diferenciação entre homonímia e polissemia é um dos mais importantes pontos teóricos ainda não devidamente compreendidos em gramática; necessitamos de pesquisa que a esclareça, com urgência.

Quadros

QUADRO 7-A: A NOÇÃO DE "CORRESPONDÊNCIA" NA LITERATURA GRAMATICAL

A noção de "correspondência", ou algo equivalente, é utilizada pela maioria dos sintaticistas. Em geral, não se encontra a preocupação de defini-la com rigor (a grande exceção é Harris; ver adiante), mas seu uso é constante nas gramáticas tradicionais, assim como em abordagens mais modernas. Naturalmente, a nomenclatura varia: os termos mais utilizados são "transformação" e "correspondência". Nesta *Gramática* optei pelo segundo termo, porque "transformação" tem uma definição bastante diferente nos modelos que admitem vários níveis de análise (a chamada "gramática transformacional", por exemplo).

A noção não está totalmente esclarecida; há problemas teóricos bastante sérios a resolver, mas aqui não será possível discuti-los. Procurei definir a noção da maneira mais explícita possível (ver a seção 7.1.) e utilizei-a na análise de maneira crucial. Se essa decisão acarreta riscos, são riscos que esta *Gramática* compartilha com a maioria das descrições atuais.

1. Na gramática tradicional

Parece ser tacitamente entendido entre os gramáticos tradicionais que existe uma relação particular entre certas estruturas; em especial, encontra-se essa relação mencionada quando se trata de passivas e ativas. Às vezes usa-se uma terminologia transformacional *avant la lettre*, como nas seguintes passagens:

> Ao passarmos a proposição para a passiva, não devemos deslocar as palavras [...]. [Maciel, 1918, p. 368]

> Sendo [o agente da passiva] o verdadeiro agente, isto é, aquele que exerce a ação, podemos transformar a construção passiva em ativa, e, neste caso, ele figurará como sujeito, passando o sujeito passivo a objeto direto [...]. [Lima, 1964, p. 245]

Os dois exemplos acima são particularmente interessantes por serem certamente anteriores às propostas transformacionalistas de Harris e de Chomsky, ou pelo menos à sua difusão no Brasil (a gramática de Rocha Lima estava, em 1964, em sua décima edição). Alguns trabalhos mais recentes, como o de Cunha & Cintra, 1985, revelam influência das idéias transformacionalistas, pelo menos na ênfase posta na relação entre estruturas, assim como no uso de árvores para mostrar a hierarquia dos constituintes.

Evidentemente, como seria de esperar, essas ocorrências de uma noção próxima da de "correspondência" na gramática tradicional em geral deixam de distinguir claramente o aspecto semântico do formal. Mas isso nem sempre ocorre; muitos autores descrevem a relação formal, como na passagem seguinte:

> Pode-se mudar a voz ativa na passiva sem alterar substancialmente o sentido da frase [...] Observe que o objeto direto será o sujeito da passiva, o sujeito da ativa passará a agente da passiva e o verbo ativo revestirá a forma passiva, conservando o mesmo tempo. [Cegalla, 1987, p. 187]

Nesse trecho, os dois aspectos estão razoavelmente bem distintos. Fica bem claro que a relação passiva/ativa não é uma relação puramente semântica: trata-se, igualmente, de uma relação formal entre duas estruturas. Uma expressão bem clara dessa distinção está numa passagem de Cláudio Brandão:

> A distinção clássica das vozes baseia-se num critério um tanto frágil: é mais formal que substancial. Nem sempre se abalizam com nitidez os discrimes entre elas. A caracterização morfológica desta ou daquela voz não coincide muitas vezes com a realidade semântica. [apud Hauy, 1983, p. 144]

Parece-me, portanto, que não é forçar o sentido do texto de muitos autores tradicionais afirmar que eles admitem uma noção muito próxima, senão idêntica, à de "correspondência", embora nunca a explicitem plenamente.

2. Na lingüística moderna

Os lingüistas que se colocam nas várias linhas coletivamente conhecidas como "estruturalismo" tendem a não reconhecer relações formais entre estruturas. Assim, Martinet parece negar a possibilidade de uma relação entre passivas e ativas, identificando as construções passivas com as construções com verbo *ser*; assim, fala ele da

> impossibilidade na qual nos encontramos de distinguir formalmente o passivo das combinações da cópula com particípios perfeitos simples [...]. [1979, p. 131]

Martinet analisa as passivas independentemente das ativas, considerando a existência de um "monema [i.e., morfema] **passivo**", composto dos diversos elementos formais

que caracterizam as orações passivas: *ser, -do, por.*

Tendo a crer que os estruturalistas mais estritos seriam levados a essa posição, por força dos pressupostos da teoria. Palmer, 1964, reconhece a necessidade de estabelecer relacionamento formal entre passivas e ativas, mas acha que isso só se pode fazer dentro de um quadro transformacional ou tagmêmico. E Robins parece confirmar a análise independente das passivas:

> Nos constituintes imediatos a sentença [*the war was started by Hitler*] pode ser simplesmente vista como o resultado de expansões sucessivas de uma estrutura sentencial básica representada por *John worked* ou *war began.* [1964, p. 242]

Ainda assim, os estruturalistas mais coerentes sentem-se obrigados a reconhecer a relação passiva/ativa quando enfrentam a tarefa de descrever línguas particulares. Bloomfield, tratando do ilocano (uma língua das Filipinas), afirma que

> Às passivas geralmente correspondem formas não-passivas cujo sujeito corresponde ao agente da passiva; nestas, uma expressão oblíqua de objeto, que em geral denota uma coisa indefinida, corresponde ao sujeito da passiva [...]. [1942, p. 98]

Temos aqui, além da noção, o próprio termo **correspondência** — aliás retomado por Greenbaum, (1969), em seu estudo sobre os advérbios.

A corrente gerativista não só reconhece as relações entre estruturas, mas é largamente baseada nelas. Nem é preciso dar exemplos; em todas as variedades da gramática gerativa a relação entre estruturas é considerada fundamental — varia apenas a maneira de descrever essa relação e de inseri-la na teoria geral.

É necessário, entretanto, colocar precisamente a noção de "correspondência" definida no capítulo 7 desta *Gramática* em relação à noção de "transformação" presente em muitos modelos gerativos. Isso será feito muito sumariamente, porque uma discussão completa do assunto nos levaria demasiadamente longe de nossos objetivos.

A correspondência se aproxima mais da transformação tal como concebida por Harris, 1957, do que da de Chomsky, 1965. Harris entende a transformação como relação sistemática entre estruturas superficiais, ao passo que Chomsky (no trabalho mencionado) define transformação como uma operação que relaciona uma estrutura com outra que lhe é subjacente, dentro de um modelo que admite diversos níveis de análise. Assim, para Chomsky, o conjunto das transformações da língua tem como função relacionar as estruturas superficiais com as estruturas profundas (ver exposição da teoria em Chomsky, 1965; ou, com exemplificação em português, em

Perini, 1976). No presente trabalho, conforme já fiz notar, restrinjo a análise à estrutura superficial, sem fazer disso uma opção teórica exclusiva. Assim, a correspondência não pode ser identificada com as transformações de Chomsky, 1965.

A fonte da noção de correspondência tal como usada nesta descrição está nos trabalhos de Harris, em particular em seu importante artigo "Co-occurrence and transformation in linguistic structure" (Harris, 1957).

Devo esclarecer, no entanto, que a posição de Harris frente aos objetivos da descrição lingüística não se confunde com a minha, nem com a atitude mais corrente em lingüística atual. Harris, em harmonia com a maior parte dos lingüistas da época, aceitava como válida a procura de "procedimentos de descoberta", isto é, um sistema de princípios que levariam, mais ou menos automaticamente, dos dados à melhor análise. Isso é negado hoje pela maioria dos lingüistas, que preferem admitir apenas a possibilidade de "procedimentos de avaliação", que possibilitam comparar duas ou mais análises e decidir qual delas é a mais adequada. As análises, em si, precisam ser fornecidas de antemão, e não há procedimentos automáticos para sua formulação: esta depende de fatores tais como a imaginação do lingüista, sua inspiração do momento e seu grau de informação sobre os fatos já conhecidos. A objetividade necessária a qualquer teoria científica só entra em jogo no segundo momento, o da avaliação das diversas teorias disponíveis.

Por outro lado, creio que Harris (pelo menos naquela época) negaria a validade de estabelecer estruturas abstratas subjacentes, das quais se derivariam as sentenças observáveis, mas que, elas próprias, nunca se realizariam. Para Harris, uma estrutura transformacionalmente relacionada a outra é também uma sentença que ocorre na língua, não uma construção abstrata. Assim, uma oração ativa é transformacionalmente relacionada a uma passiva, embora ambas sejam superficialmente realizadas e, portanto, nenhuma delas seja "subjacente" no sentido de Chomsky. Por minha parte, tendo a acreditar que há generalizações válidas em sintaxe que exigem a postulação de estruturas abstratas, nunca realizadas; acho que a literatura gerativa vem mostrando isso com certa segurança.

Mas a noção de relacionamento superficial entre estruturas (incluída igualmente nos modelos chomskianos, embora de maneira indireta) me parece necessária. Concordo com a afirmação de Michael Kac de que

> O pressuposto harrisiano de que as transformações são relações de sentença para sentença é um pressuposto significativo, cujo valor e interesse têm sido injustamente negados pelos gramáticos gerativistas. [1978, p. 470]

Harris define as transformações de maneira mais liberal do que se concebem neste trabalho: ele inclui como casos de relação transformacional grupos de estruturas que não seriam correspondentes segundo minha análise. Entretanto, é necessário reconhecer que as restrições que imponho à correspondência não estão suficientemente fundamentadas, e é possível que venham a ser reformuladas no sentido de uma aproximação do trabalho de Harris. De qualquer forma, minha dívida para com as propostas de Harris é crucial e precisa ser registrada aqui. Em uma palavra, meu trabalho em sintaxe é em grande parte uma derivação da análise praticada por Harris.

A maior diferença é que eu gostaria de limitar a noção de "correspondência", assim como muitas outras entidades propostas neste estudo, ao papel de instrumento de descrição de regularidades superficiais. Note-se que essa posição permite continuar aceitando a possível existência de regularidades importantes não formuláveis em termos estritamente superficiais.

QUADRO 7-B: FUNÇÕES SINTÁTICAS EM ESTRUTURAS CORRESPONDENTES

Vimos na seção 7.1.3. que quando duas (ou mais) orações são totalmente correspondentes os termos correlatos têm sempre a mesma função. Assim, em

(1) a. Leo comeu o peixinho.
  b. O peixinho, Leo comeu.

o termo *Leo* é sempre sujeito, *o peixinho* é sempre objeto direto e *comeu* é sempre núcleo do predicado. Neste quadro, vamos expor as razões que justificam a análise paralela das estruturas totalmente correspondentes.

A alternativa, naturalmente, seria atribuir uma função nova a *o peixinho*, justificando-se essa nova função, no caso, pela posição diferente que o termo ocupa em (1b). Seria necessário, para isso, modificar a definição de objeto direto, retirando-se o traço [+ Ant], que especifica justamente a possibilidade de ocorrer anteposto. Por que essa solução não pode ser adotada?

A primeira razão tem a ver com a simplicidade da análise. Se *o peixinho* em (1b) não for OD, terá de ter uma função inteiramente nova, pois não se encaixa em nenhuma das funções já definidas. Aqui teremos uma função a mais, o que pode não parecer um grande problema; mas a situação se repete nos outros casos de correspondência total. Veja-se, por exemplo, o caso do posicionamento alternativo do adjunto oracional:

(2) a. Indubitavelmente, o prefeito estava envolvido na negociata.
  b. O prefeito, indubitavelmente, estava envolvido na negociata.

c. O prefeito estava, indubitavelmente, envolvido na negociata.

d. O prefeito estava envolvido, indubitavelmente, na negociata.

e. O prefeito estava envolvido na negociata, indubitavelmente.

Essas frases são totalmente correspondentes; segundo nossa análise atual, todas elas incluem um adjunto oracional. Se não mantivermos o princípio de que estruturas correspondentes encerram as mesmas funções, porém, será necessário distinguir aí cinco funções, cuja única diferença está na posição que o termo ocupa na oração. Como se vê, a análise adotada na *Gramática*, em que *indubitavelmente* é AO nas cinco frases, é bem mais simples.

Esse argumento, por si só, não pode ser decisivo: afinal de contas, há ocasiões em que temos de distinguir funções diferentes. Mas há outras razões para manter a mesma análise em estruturas totalmente correspondentes.

Nossa análise permite capturar certas generalizações acerca da transitividade dos verbos. Voltando aos exemplos (1), sabemos que o verbo *comer* admite objeto direto. Se analisássemos *o peixinho* em (1b) como outra coisa qualquer, teríamos uma frase sem objeto direto em

(1) b. O peixinho, Leo comeu.

Logo, deveria ser possível acrescentar um OD à frase (pois o verbo o permite). Mas não é:

(3) * O peixinho, Leo comeu um pastel.

Dentro dessa análise, a inaceitabilidade de (3) é inteiramente misteriosa. Já se aceitarmos que *o peixinho* em (1b) é OD (topicalizado), a inaceitabilidade de (3) derivará automaticamente do princípio, válido para toda a língua, de que nenhuma oração pode ter mais de um objeto direto.

Essas são as razões sintáticas para se manter o princípio de que estruturas totalmente correspondentes incluem as mesmas funções sintáticas, ocupadas estas pelos mesmos itens léxicos. Se *o peixinho* é OD em uma frase F, será OD em toda e qualquer frase totalmente correspondente a F.

Passemos agora a um segundo problema: será que o raciocínio acima aplicado pode ser transferido para o caso da correspondência parcial? A resposta, como já adiantei na seção 7.1.3., é negativa. Vejamos agora as razões para isso.

A razão principal é que o segundo argumento levantado em favor da identidade de funções para os casos de correspondência total — isto é, a possibilidade de exprimir a transitividade dos verbos da maneira mais compacta — não se aplica nos casos de correspondência parcial.

Ao considerarmos os casos de correspondência total, vimos que um OD topicalizado tem exatamen-

te o mesmo efeito de um OD não-topicalizado. Mas a situação é diferente com a correspondência parcial. Suponhamos que atribuíssemos a mesma função ao objeto da ativa e ao sujeito da passiva. Acontece que há verbos que admitem o primeiro, mas não o segundo (por não admitirem passiva). É o caso de *morrer, ter, dormir* e vários outros verbos. Esse fato nos obriga a especificar, para cada verbo que admite OD, se é ou não possível formar uma passiva (transformando o OD em sujeito, entre outras coisas). As condições de ocorrência de objetos diretos e de sujeitos de passivas são diferentes, e é necessário expressá-las separadamente para cada verbo. Ou seja, nas frases

(4) a. Leo comeu o peixinho.

b. O peixinho foi comido por Leo.

não podemos atribuir a mesma função a *o peixinho* nas duas frases.

Se analisarmos as passivas e as ativas da mesma maneira (isto é, com as mesmas funções para os mesmos sintagmas, ainda que estes estejam ordenados diferentemente), já não será possível afirmar que *dormir* aceita OD (como em *ele está dormindo um sono tranqüilo*) — em certos casos aceitaria OD, em outros o recusaria. Como se viu no capítulo 6, essa possibilidade de recusa de complemento em certas construções não se verifica com nenhum verbo: um verbo tem traços de transitividade independentemente da construção em que aparece.

Outro argumento é que, se mantivéssemos as mesmas funções para cada constituinte oracional nos pares ativa/passiva, teríamos certamente que refazer as bases da análise sintática que estamos adotando. Isto é, se *o peixinho* for OD tanto em (4a) quanto em (4b), já não será possível dizer que o NdP concorda com o sujeito (pois estaria concordando com o OD em (4b)). A força desse argumento é que (muito provavelmente) essa reformulação levaria a uma maior complicação geral da análise.

Em conclusão: em um grupo de estruturas que sejam totalmente correspondentes, cada sintagma particular tem sempre a mesma função sintática em todas as estruturas em questão. Mas se a correspondência for parcial pode haver mudanças de função do mesmo sintagma, ao se passar de uma estrutura para outra.

QUADRO 10-A: AGENTE E INTENCIONALIDADE

Na seção 10.1., defini o agente como "a entidade que provoca a ação denotada pelo verbo" e acrescentei que se entende o agente como dotado de intencionalidade. Ou seja, o agente será sempre um ser animado, no mais das vezes humano. E em 10.2.7. utilizei o teste do acréscimo de um adverbial orientado para o agente (AOPA), que serviria para detectar a presença de um agente

na representação semântica de uma sentença. Esse teste permite afirmar que a frase

(1) A janela foi quebrada.

possui agente em sua representação semântica, ao contrário de

(2) A janela quebrou.

que não possui agente. Sabemos disso porque um adverbial como *propositalmente* (um AOPA) pode ser acrescentado a (1), mas não a (2).

Neste quadro pretendo discutir um pouco mais detalhadamente a validade desse teste, assim como a do conceito de "agente" formulado na seção 10.1.

Segundo Jackendoff, o agente não precisa ser dotado de intencionalidade (ou volição própria); para ele, existiria um agente na frase

(3) A ventania atirou Bilico na piscina.

Afirma ele que

> [...] o agente não está necessariamente agindo voluntariamente [...] a volição ou intencionalidade é uma propriedade opcional de um **ator** [*actor*], e não precisa ser representada adicionalmente comoparte da função CAUSA. [Jackendoff, p. 176]

Não pretendo discutir aqui a adequação da idéia de Jackendoff. Antes, tentarei mostrar que, mesmo que ele tenha razão, o teste dos AOPAs é válido para o que nos interessa.

O problema, evidentemente, é que o acréscimo de um AOPA a uma frase pode gerar inaceitabilidade mesmo se houver agente na representação semântica — basta que esse agente seja inanimado, e portanto destituído de volição. Por isso não se pode dizer

(4) * A ventania atirou Bilico na piscina de propósito.

Mas já de saída pode-se salvar alguma coisa do teste: a possibilidade de acrescentar um AOPA sem produzir inaceitabilidade é evidência da presença de agente (no caso, animado) na representação semântica. Desse modo, a diferença apontada entre (1) e (2) se mantém, na medida em que (1) necessariamente possui um agente, ao passo que (2) ou não possui ou então possui agente inanimado. O problema, pois, se reduz a tentar mostrar que (2) não tem agente de espécie nenhuma em sua representação semântica.

Vamos então centrar nossa discussão no exemplo

(5) * A janela quebrou de propósito.

Como vimos, parece haver duas explicações para a inaceitabilidade dessa frase: ou a ausência de agente em sua representação semântica, ou a presença de um agente, porém inanimado.

Entretanto, há indicações suficientes de que frases como essa não nos obrigam a entender um agente inanimado (o que seria o caso se tal

agente estivesse presente na representação semântica). A primeira indicação se baseia na intuição dos falantes, que ao ouvirem

(2) A janela quebrou.

não são obrigados a entender que o evento foi provocado por um agente inanimado. Pode-se perfeitamente entender que alguém quebrou a janela e utilizar (2) por alguma razão — por exemplo, porque a expressão do agente seria irrelevante. Além disso, podem-se inserir tais frases em um contexto que deixe claro que o agente é animado, e o resultado é perfeitamente aceitável:

(6) Bilico forçou até que a janela quebrou.

Ora, se *a janela quebrou* tivesse agente inanimado na representação semântica, haveria uma contradição em (6). Como essa contradição não existe (haja vista a aceitabilidade da frase), temos de concluir que:

(a) *a janela quebrou* tem agente animado na representação semântica;

ou então que

(b) *a janela quebrou* não tem nenhuma especificação de agente na representação semântica (o que deixaria em aberto a possibilidade de haver um agente, e este ser animado).

Mas acontece que a alternativa (a) está em contradição com a inaceitabilidade de

(5) * A janela quebrou de propósito.

Se a frase tivesse agente animado, o acréscimo do AOPA seria possível; como não é, precisamos admitir a alternativa (b) como a verdadeira.

Em resumo, podemos manter a afirmação de que em (1) a representação semântica inclui um agente (animado), mas em (2) não inclui agente nenhum.

QUADRO 10-B: SOBRE A SEMÂNTICA DO PREDETERMINANTE

1. Papel da interpretação distributiva

Na seção 4.3., estudamos as possibilidades de colocação do predeterminante (PDet) dentro da oração. Ali foi proposto que a restrição principal é que o PDet precisa posicionar-se entre quaisquer termos de nível oracional. Por exemplo, podemos ter

(1) a. Todos os móveis estão sendo vendidos.

b. Os móveis estão todos sendo vendidos.

c. Os móveis estão sendo todos vendidos.

Essa regra sofre uma exceção, que é a possibilidade de ocorrência dentro do SN, logo antes do modificador externo, como em

(2) Os motoristas todos do Brasil entraram em greve.

No presente quadro não levaremos em conta essa exceção; vamos apenas examinar algumas restrições à colocação do PDet entre os termos de nível oracional, com a idéia de que tais restrições não são sintáticas, mas decorrem de princípios de boa formação semântica.

Devo esclarecer que, no que pesem alguns bons trabalhos já realizados na área, as condições que governam o posicionamento do PDet ainda apresentam problemas. Aqui estou apresentando uma sistematização inicial (mais que propriamente uma análise) válida para boa parte dos dados. Mas ainda há o que esclarecer a respeito.

Vejamos primeiramente alguns exemplos que levantam problemas para a regra de posicionamento do PDet. Podemos começar com

(3) * Os fatos surpreenderam todos Raul.

Esse exemplo já foi estudado por Simões (1974); mas a solução que darei aqui é bem diferente da proposta por essa autora.

Em (3), o PDet, *todos*, está colocado entre o NdP e o objeto direto, que são constituintes de nível oracional; a frase, portanto, é sintaticamente bem formada e deveria ser aceitável, mas não é. Tentarei mostrar que ela apresenta uma má formação semântica, e por isso é inaceitável (evidentemente, uma frase só é aceitável quando é bem formada segundo todos os componentes da gramática: fonologia, morfologia, sintaxe, semântica).

Para ver o que há de errado semanticamente com (3), comparemo-la com

(4) Os rapazes conseguiram todos um bom emprego.

Essa frase tem estrutura sintática basicamente semelhante à de (3): um sujeito no plural, um NdP, um PDet posposto (que se deve relacionar semanticamente com o sujeito) e um objeto direto. No entanto, (4) é perfeitamente aceitável, ao passo que (3) não é. A pesquisa feita até o momento não conseguiu explicar essa diferença de aceitabilidade com base na estrutura sintática.

No entanto, há uma diferença semântica que, como veremos, é provavelmente responsável pela estranheza de (3). Em (4), entende-se que cada rapaz conseguiu um bom emprego; ou seja, embora esteja no singular, *um bom emprego* se deve entender como referindo-se a vários empregos (tantos quantos forem os rapazes). Esse tipo de interpretação se chama **interpretação distributiva**. Vemos outro exemplo em frases como

(5) Essas casas vão ganhar uma pintura nova.

Aqui também a acepção mais imediata é a de que cada casa receberá uma pintura.

Ora, observa-se que frases com a estrutura exemplificada em (3) e (4) são aceitáveis quando é possível atribuir-lhes uma interpre-

tação distributiva, e inaceitáveis em caso contrário. Em (3), a interpretação distributiva é impossível, porque não há meio de distribuir Raul pelo conjunto dos fatos ("um fato para cada Raul", ou vice-versa, não faz sentido).

Essa restrição é bastante misteriosa (afinal, que fator ligaria a posição posposta do PDet com a interpretação distributiva?), mas parece que existe realmente. Quando o PDet está antes do NdP, a interpretação distributiva pode aparecer, mas não é obrigatória:

(6) Todos os fatos surpreenderam Raul.

(7) Todos os rapazes conseguiram um bom emprego.

Somente em (7) a interpretação é distributiva; não obstante, isso não atrapalha a aceitabilidade de (6), porque a interpretação não-distributiva também é possível.

Como se vê, deve haver vários fatores governando o aparecimento desse tipo de interpretação. O que nos interessa em especial pode ser formulado como uma regra semântica, a saber:

1ª Regra de Interpretação Distributiva (ID)

A cadeia [Sujeito$_i$ ... PDet$_i$ + OD] *recebe interpretação distributiva.*

[Entende-se que os elementos subscritados com "$_i$" são correferentes.]

(Chamei a regra de "1ª Regra" porque certamente há outras regras semânticas que igualmente atribuem interpretação distributiva, ou ID, em outras circunstâncias.)

Explicando: falei de "cadeia" porque a interpretação distributiva, por sua própria natureza, só se aplica a conjuntos não-unitários de elementos. No caso, ela estabelece um mapeamento um-a-um entre o sujeito e o objeto direto, de modo que, a cada elemento do sujeito, corresponderá um elemento do objeto: em (4), a cada rapaz corresponde um bom emprego. Assim, não faria sentido dizer que o objeto direto, sozinho, recebe interpretação distributiva; antes, devemos dizer que a cadeia **sujeito** (com seu **PDet** correferente) **— objeto direto** é que recebe a interpretação distributiva.

É muito curioso observar que essa regra só se aplica obrigatoriamente na presença de um objeto direto. Qualquer outro termo que se siga a um PDet posposto fica isento; assim, pode-se dizer

(8) As garotas se declararam todas a Giacomo.

muito embora não seja possível distribuir "cada garota a um Giacomo". Aqui, como não há objeto direto (*a Giacomo* é adjunto circunstancial), a interpretação distributiva não precisa surgir; ao que parece, surge opcionalmente, como quando se coloca *a um rapaz* no lugar de *a Giacomo.*

Essa é pois a razão da diferença de aceitabilidade entre as frases (3) e (4). Em ambas, a 1ª Regra de ID atribui, obrigatoriamente, interpretação distributiva à cadeia formada por **sujeito + PDet + objeto direto**. Mas em (3) isso resulta em mapeamento um-a-um entre um conjunto de elementos (*os fatos*) e um elemento necessariamente único (*Raul*); como resultado, a interpretação semântica final da sentença é anômala. Em (4) isso não acontece, porque não há problemas em fazer o mapeamento de "um bom emprego para cada rapaz".

2. Semântica do PDet posposto

A frase seguinte parece, à primeira vista, trazer um problema para a análise da seção precedente:

(9) A renúncia surpreendeu todos os deputados.

A formulação da 1ª Regra de ID parece exigir a atribuição de interpretação distributiva neste caso; mas é evidente que isso não acontece, porque (9) não inclui em seu significado a idéia de "uma renúncia para cada deputado". Assim, a frase não é anômala porque não contém interpretação distributiva. Mas por que não contém?

A diferença relevante entre (9) e

(4) Os rapazes conseguiram todos um bom emprego.

é que em (4) *todos* se refere ao sujeito (*os rapazes*), e em (9) *todos* se refere ao objeto direto (*os deputados*). Isso está indicado na formulação da 1ª Regra de ID pelos pequenos "ii" subscritos ao sujeito e ao PDet: a regra só funciona quando o PDet se refere ao sujeito.

Nosso problema, portanto, passa a ser: Por que é que, em (4), o PDet se refere ao sujeito, e em (9) se refere ao objeto direto? Vamos agora abordar esse problema e veremos que o sistema de atribuição da referência do PDet interage com a 1ª Regra de ID para explicar uma variedade bastante ampla de sentenças da língua.

Precisamos de uma regra semântica que relacione o PDet com um SN determinado dentro da oração. Sabemos que em (4) o PDet se relaciona com o SN *os rapazes*, embora esteja na verdade contíguo ao SN *um bom emprego*. Já em (9) o PDet se relaciona com o SN que imediatamente se segue, *os deputados*; a regra deverá dar conta dessas diferenças, entre outras.

A regra é a seguinte:

Regra de indexação do PDet (RIP)

O PDet se relaciona semanticamente (é coindexado) com:

(a) o SN com o qual está em construção, formando com ele um sintagma;

(b) o SN contíguo à esquerda;

(c) um SN topicalizado;

(d) o 1º SN à esquerda.

(As quatro partes da regra, (a)–(d), se aplicam na ordem, e a aplicação de uma exclui a aplicação das outras; ou seja, apenas uma das partes se pode aplicar em cada caso.)

A melhor maneira de explicar a RIP é tomar uma série de exemplos e trabalhá-los; nos casos relevantes, aplicaremos também a 1ª Regra de ID para atribuir, se for o caso, a interpretação distributiva. Finalmente, em certos casos poderá resultar uma anomalia semântica, o que marcará a frase como inaceitável.

Vamos começar com

(6) Todos os fatos surpreenderam Raul.

O PDet a ser relacionado é *todos*. Ele está em construção com o SN *os fatos*, pois se coloca imediatamente antes dele e não desobedece a nenhuma condição sintática (por exemplo, o SN contém Det, o que é uma exigência do PDet). Assim, (a) estabelece que *todos* e *os fatos* devem ser co-indexados, ou seja, entendidos como referindo-se à mesma entidade.

Por outro lado, a 1ª Regra de ID não se aplica aqui, pois não temos um PDet imediatamente antes de um objeto direto. Como resultado, *todos* se relaciona com *os fatos*, e não há razão semântica para excluir a frase, que é aceitável.

Vejamos agora

(3) * Os fatos surpreenderam todos Raul.

Já examinamos essa frase na seção precedente; agora vamos revê-la, incluindo a RIP. O PDet vem logo antes do SN *Raul*, mas não está em construção com ele (*todos Raul* não forma um sintagma); se estivesse, o sintagma seria mal formado sintaticamente (o SN não contém Det) e semanticamente (*todos* é plural, e *Raul* é singular). Dessa forma, (a) não se aplica.

Passamos a (b), que tampouco se aplica, pois não há nenhum SN imediatamente à esquerda do PDet. E (c) também deixa de se aplicar, porque não há SN topicalizado.

Resta (d), que relacionará *todos* ao primeiro SN à esquerda, que é *os fatos*. Mas neste caso, como já sabemos, a 1ª Regra de ID tem condições de se aplicar, impondo interpretação distributiva à cadeia [*Os fatos* + *todos* + *Raul*]. E sabemos que isso acarreta uma anomalia semântica (ver a seção precedente deste quadro); por isso, (3) é inaceitável.

O caso de

(4) Os rapazes conseguiram todos um bom emprego.

é muito semelhante ao de (3): *todos* se relaciona com o sujeito *os rapazes*, pela aplicação de (d); e a interpretação distributiva também ocorre. Apenas, não surge anomalia porque não há estranheza na idéia de "um bom emprego para cada rapaz".

Outro exemplo paralelo é o de

(10) Os elefantes escaparam todos do zoológico.

O relacionamento de *os elefantes* com *todos* se faz através de (d), como nos dois exemplos anteriores. Não se atribui interpretação distributiva, pois não há objeto direto; o resultado é uma frase aceitável, entendendo-se *todos* como referindo-se a *os elefantes.*

Vejamos agora

(11) Os bolos, os meninos comeram todos.

O problema aqui é que *todos* se refere claramente a *os bolos,* e não ao SN mais próximo à esquerda, *os meninos.* Atribuo isso ao fato de que *os bolos* está claramente topicalizado (transportado para o início da sentença). Nesses casos, parece que precisamos de uma sub-regra especial, a saber, a parte (c) da RIP. A parte (a) não se aplica, porque o PDet não está em construção com nenhum SN — ou seja, não forma um sintagma juntamente com um SN; (b) também não se aplica, porque *todos* não está contíguo a um SN. Passamos a (c), que se aplica porque há um SN topicalizado; assim, *todos* se relacionará semanticamente com esse SN topicalizado, *os bolos.* Finalmente, a interpretação distributiva não aparece, porque a frase não contém a cadeia descrita na 1ª Regra de ID.

Um novo exemplo é

(12) * Os bolos, minha sobrinha todos comeu.

Acredito que essa frase é inaceitável em português padrão usual (embora talvez pudesse aparecer como um caso extremo de hipérbato na poesia antiga). Vejamos o que causa essa inaceitabilidade.

Aplicando RIP, vemos que (a) não se aplica; (b) se aplica, relacionando *todos* com *minha sobrinha.* Mas esse relacionamento causa anomalia semântica, porque *todos* é masculino e plural, e *minha sobrinha* é feminino e singular. Daí deve provir a inaceitabilidade de (12). Se a frase fosse

(13) As panquecas, minhas sobrinhas todas comeram.

retirando-se as incompatibilidades de gênero e número, a anomalia desaparece; (a) ainda não se aplica, mas (b) relaciona, sem problemas, *todas* com *minhas sobrinhas.* Observe-se que, embora o SN *as panquecas* esteja topicalizado, o PDet não se relaciona com ele, porque é preciso aplicar (ou tentar aplicar) (b) antes de (c). E, uma vez aplicada (b), já não se pode aplicar (c) ou (d): já vimos que a aplicação de uma das partes de RIP exclui a aplicação das outras.

Finalmente, vejamos um exemplo talvez mais problemático:

(14) Os rapazes conseguiram todos os melhores empregos.

O problema potencial aqui é que *todos* pode ser entendido como fazendo parte de um sintagma *todos os melhores empregos,* e nesse caso (a) relacionará *todos* com *os melhores empregos.* Mas também se pode entender

todos como um PDet posposto, não fazendo sintagma com *os melhores empregos,* e nesse caso (d) o relacionará com *os rapazes.*

Parece, com efeito, que ambas as interpretações estão disponíveis; e uma leitura rápida pode levar o leitor a uma ou outra. Mas a ambigüidade é mais aparente que real. Na fala, a segunda interpretação (*todos + os rapazes*) só ocorre se houver uma cesura entonacional logo antes de *os rapazes*:

(15) Os rapazes conseguiram todos/um bom emprego.

e na escrita esse efeito se consegue separando *todos* com vírgulas, o que o impede de ser entendido como parte de um sintagma com *os melhores empregos*:

(16) Os rapazes conseguiram, todos, os melhores empregos.

Parece-me que (16) não é ambígua, e que **todos** se relaciona obrigatoriamente com *os rapazes* — aliás, justamente como prevê a formulação de RIP.

3. Um falso problema

Agora vamos examinar, muito brevemente, alguns dados que parecem oferecer um problema para a análise do PDet como elemento externo ao SN; esses dados parecem indicar que o PDet faz exigências de compatibilidade mesmo quando está à direita do SN, e não contíguo a ele. Por exemplo, sabemos que (no português padrão) a presença de um PDet logo antes de um SN acarreta a exigência de que o SN tenha um determinante:

(17) Todos os elefantes escaparam do zoológico.

(18) * Todos elefantes escaparam do zoológico.

Isso pode acontecer porque em tais casos *todos* está em construção com *os elefantes* (isto é, pertencem ao mesmo sintagma). A idéia é que as exigências de compatibilidade só funcionam dentro de um constituinte.

No entanto, as frases abaixo parecem mostrar o contrário:

(19) Os elefantes escaparam todos do zoológico.

(20) * Elefantes escaparam todos do zoológico.

Aqui parece que *todos,* ainda que posicionado à direita do SN *(os) elefantes,* e portanto não em construção com ele, continua impondo exigências de compatibilidade: o determinante continua sendo obrigatório nesses casos. Será que teremos de admitir que as compatibilidades funcionam também "a distância"?

A resposta é que não. Vou mostrar que o que exclui (20) não é uma exigência de compatibilidade originada no PDet, mas um choque semântico de natureza totalmente diferente. Para isso, observemos primeiro dois fatos:

(a) quando o verbo está no pretérito perfeito, a interpretação do sujeito é necessariamente específica (e referencial; ver 10.4.).

Isso quer dizer que não podemos ter um sujeito de significado genérico com o verbo no perfeito. Assim, por exemplo, em

(21) O elefante come mais de 250 empadas por dia.

o elefante pode ser entendido como significando "todo e qualquer elefante" (genérico); (21) será, nesse caso, uma afirmação acerca dos elefantes em geral. Mas se pusermos o verbo no perfeito,

(22) O elefante comeu mais de 250 empadas por dia.

essa interpretação já não será possível. Só se pode entender (22) como uma afirmação sobre um elefante específico, e não como uma verdade geral acerca de elefantes.

O segundo fato é

(b) um SN composto unicamente de um nome comum no plural só se pode interpretar genericamente.

Assim, o SN *elefantes* só se entende como referindo-se a "todo e qualquer elefante", como em

(23) Elefantes comem bastante.

Não seria possível interpretar (23) como afirmando algo a respeito de vários elefantes específicos.

Voltemos agora a nossos exemplos:

(19) Os elefantes escaparam todos do zoológico.

(20) * Elefantes escaparam todos do zoológico.

Agora podemos explicar a inaceitabilidade de (20) a partir de uma anomalia semântica: o sujeito *elefantes*, sendo composto apenas de um nome comum no plural, deve ter interpretação genérica; mas o verbo está no perfeito, o que nos obriga a interpretar o sujeito como específico (isto é, não-genérico). A contradição resulta em anomalia semântica, e portanto inaceitabilidade.

Note-se que essa explicação supõe que a presença do PDet não tem nada a ver com a aceitabilidade ou inaceitabilidade dessas frases. E, de fato, se retirarmos o PDet, os resultados serão exatamente os mesmos:

(24) Os elefantes escaparam do zoológico.

(25) * Elefantes escaparam do zoológico.

Outra conseqüência da explicação dada é que se o verbo for colocado em um tempo que admite interpretação genérica a inaceitabilidade de (25) deverá desaparecer; isso é confirmado:

(26) Elefantes escapam do zoológico (todos os dias).

E, naturalmente, mesmo se acrescentarmos um PDet, o resultado não pode variar, o que é igualmente confirmado:

(27) Elefantes sempre escapam todos do zoológico.

ou, para termos um exemplo mais plausível,

(28) Elefantes morrem todos antes dos 40 anos.

Em conclusão: a inaceitabilidade de frases como (20) não tem nada a ver com presumíveis exigências de compatibilidade originadas no PDet; e pode-se manter o princípio de que as compatibilidades só funcionam dentro de um constituinte.

QUADRO 10-C: ORAÇÕES SEM SUJEITO: SINTAXE E SEMÂNTICA

1. Sintaxe

Vimos na seção 3.2.2.3. que a definição de sujeito aqui adotada nos força a considerar como sem sujeito frases como

(1) Cheguei ontem de Goiânia.

Essa análise costuma encontrar alguma resistência, porque as pessoas estão muito acostumadas a ver em (1) um "sujeito oculto", que seria responsável tanto pela concordância quanto por certos traços da interpretação semântica. Neste quadro, vou mostrar que há indicações razoavelmente claras de que (1) não tem, com efeito, sujeito. Na seção 2 deste quadro retomo o problema, mostrando que igualmente na semântica não é possível tratar paralelamente os dois tipos de sujeito. A conclusão é que não vale a pena postular sujeitos ocultos.

Em primeiro lugar, recordarei que se (1) tem sujeito, será preciso trocar a definição de sujeito como termo que está em relação de concordância com o verbo. Essa definição se refere a um termo **explícito**, e é claro que não há nenhum elemento explícito em (1) que esteja em relação de concordância com o verbo. Como não é essa minha opção, deixarei a quem pretenda defender a existência de sujeito em (1) o trabalho de formular uma nova definição.

Antes de examinar o argumento, observarei que o interesse de se postular os "sujeitos ocultos" vem da presumida possibilidade de tratá-los da mesma maneira que os sujeitos claros. Assim, se (1) tiver um sujeito (oculto), poderemos explicar a concordância da mesma maneira que em

(2) Eu cheguei ontem de Goiânia.

Ou seja, o verbo está concordando com um elemento representado por *eu* e que está explicitado em (2) e implícito em (1). A regra de concordância trata da mesma forma o sujeito claro e o oculto.

Igualmente, a interpretação semântica pode ser feita paralelamente. Tanto em (1) quanto em (2) entende-se que o agente de *chegar* é "eu"; isso se explica porque, em ambos os casos, o sujeito é *eu*, e sujeitos claros são interpretados da mesma maneira que sujeitos ocultos.

Deixemos pois estabelecido que a motivação essencial para se postular o sujeito oculto é o interesse em tratá-lo, na semântica e na sintaxe, como um sujeito normal. A única diferença ficaria na explicitação de um e não do outro, na forma final da sentença. Neste quadro, mostrarei que isso não pode ser sustentado — isto é, em certos pontos da sintaxe somos forçados a tratar o sujeito claro e o oculto diferentemente; isso constitui um argumento contra a existência de sujeitos ocultos, pois prejudica sua motivação básica.

Vejamos pois as razões sintáticas para não aceitar o sujeito oculto em (1). O argumento provém de certas características dos verbos *parecer* e *dever*. Esses verbos, como a maioria dos verbos, fazem exigências quanto à forma de seus complementos; assim, *dever* só aceita complemento no infinitivo, e *parecer* aceita complemento no infinitivo ou no indicativo (ver a seção 6.4. para detalhes desse tipo de exigência):

(3) Vítor deve chegar amanhã.

(4) * Vítor deve que chega/chegue amanhã.

(5) Vítor parece estar gripado.

(6) Vítor parece que está gripado.

A exigência que nos interessa aqui é que o complemento de *dever* e *parecer* **não pode ter sujeito**. Isso se aplica seja qual for o modo da subordinada, como se vê nos exemplos

(7) * Vítor parece Chiquinho estar gripado.

(8) * Vítor parece que Chiquinho está gripado.

(9) * Vítor parece que ele está gripado.

Note-se como (9) é inaceitável mesmo se entendermos *ele* como referindo-se ao mesmo Vítor. Se houver uma cesura ou pausa (marcada por vírgula) após *Vítor*, a frase fica aceitável:

(10) Vítor, parece que ele está gripado.

Mas nesse caso a estrutura da frase é mudada, e *Vítor* já não é o sujeito; é um tópico, elemento externo à oração. Isso fica evidente se tivermos um SN plural no lugar de *Vítor*; nesse caso, os exemplos seriam:

(11) Os alunos parecem que gostam de Çãozinha.

(12) * Os alunos parecem que eles gostam de Çãozinha.

(13) Os alunos, parece que eles gostam de Çãozinha.

Observe-se a falta de concordância de *parecer* com *os alunos* em (13), que se explica justamente por *os alunos* ser um elemento externo à oração.

Assim, é seguro concluir que *parecer* sofre a restrição de que seu complemento não tenha sujeito; e é fácil mostrar que o mesmo se dá com *dever*.

Ora, podemos retirar desses fatos um argumento em favor da hipótese de que

(1) Cheguei ontem de Goiânia.

não tem sujeito. Consideremos os exemplos

(6) Vítor parece que está gripado.

(9) * Vítor parece que ele está gripado.

Se admitirmos que *está* tem sujeito em ambas as frases [oculto em (6), claro em (9)], ficará difícil exprimir a condição que exclui (9). Terá de ser alguma coisa como

> O verbo *parecer* proíbe sujeito **claro** em seu complemento.

Mas acontece que essa formulação anula a vantagem, já mencionada, de se postular o sujeito oculto: a pos-sibilidade de tratá-lo na maior parte da gramática como algo inteiramente paralelo ao sujeito claro. Aqui somos obrigados a tratar diferentemente o sujeito claro do sujeito oculto.

Na análise proposta neste livro, evidentemente, o problema nem chega a colocar-se, porque os dois casos são diferentes desde o início: *está* tem sujeito em (9) — o que acarreta sua exclusão do número das frases bem formadas —, mas não tem sujeito em (6). A condição pode ser simplesmente de que *parecer, dever* etc. proíbem sujeito em seu complemento.

2. Semântica

Agora vejamos algumas razões semânticas que apóiam a análise de (1) como oração sem sujeito.

A primeira tem a ver com o caráter redundante da informação sobre o sujeito — ou, mais exatamente, sobre a entidade semântica, qualquer que seja, expressa pelo sujeito em cada oração.

É evidente que em uma frase como

(14) Vendi meu jegue.

será preciso que uma regra semântica interprete a terminação verbal como referindo-se a determinada pessoa (no caso, o falante). A presença do falante é necessária na interpretação semântica da frase, e a única fonte possível para esse elemento é a terminação verbal; assim deve haver alguma regra cujo resultado é algo como "um verbo (no pretérito perfeito) terminado em *-i* refere-se ao falante". Se a terminação fosse *-emos*, referir-se-ia ao falante mais uma ou mais pessoas, e assim por diante. Nenhum outro elemento da frase poderia dar-nos essa informação; portanto, podemos dizer que não há redundância na interpretação do falante em (14).

Agora vejamos o exemplo

(15) Eu vendi meu jegue.

A interpretação de (15) é idêntica à de (14) — isto é, as duas frases são sinônimas. Ambas incluem, em particular, a informação de que o agente de "vender" é o falante. Essa informação provém, em (15), da terminação verbal [tal como em (14)] e também da presença do sujeito *eu*. Ou seja, a mesma informação provém de duas fontes e é, portanto, redundante. Se apenas considerássemos os interesses da comunicação, diríamos que uma das fontes é dispensável; mas, é claro, a língua tem outras exigências, de modo que *eu* pode ser dispensado, mas o sufixo verbal não.

Admitindo que a expressão do elemento ao qual o verbo se refere é redundante em (15), será possível mostrar que a análise proposta nesta *Gramática* é superior à tradicional, por capturar esse fato de maneira mais simples e natural.

Suponhamos que aceitamos a análise tradicional, que como sabemos postularia um "sujeito oculto" em (14). A vantagem de postular esse sujeito oculto é poder fazer afirmações gerais sobre sujeitos (e suas conseqüências, como a concordância verbal), sem precisar especificar (14) como um caso especial. Mas para explicitarmos o fato de que (15) é redundante será necessário justamente tratar diferentemente os sujeitos claros dos ocultos. Teremos de dizer que (14) não é redundante porque, apesar de ter sujeito, este não aparece explicitamente; em outras palavras, temos um sujeito cujo efeito semântico é suprimido em certos casos. A representação da redundância é indireta, e na verdade não terá nada a ver com a estrutura da sentença (que, afinal, inclui um sujeito), mas com um fato à parte, o caráter "oculto" desse sujeito.

Na análise proposta neste livro, por outro lado, a redundância no caso de (15) pode ser descrita de maneira mais simples. Diremos simplesmente que (15) é redundante porque a mesma informação vem de duas fontes: do sujeito e da terminação verbal. E (14) não é redundante porque não tem sujeito, apenas a terminação verbal. Essa análise, entre outras coisas, parece refletir mais de perto o desempenho do ouvinte, que certamente não reconstrói um sujeito ao interpretar (14); ao invés disso, ele interpreta diretamente a terminação verbal. Concluo que, neste particular, a análise aqui proposta é superior à análise tradicional.

QUADRO 11-A: PROBLEMAS DA ANÁLISE DA REFERÊNCIA PRONOMINAL

A regra R5, vista na seção 11.1.2., tem como conseqüência que nunca se pode ter um pronome não-reflexivo e seu antecedente dentro da mesma oração. Assim na frase

(1) Celinha a penteou.

R5 estabelece que *Celinha* não é correferente de *a*; isso está de acordo com a intuição dos falantes, que com efeito entendem *Celinha* e o pronome como referindo-se a pessoas diferentes. Se o pronome fosse reflexivo, R5 o marcaria como correferente de *Celinha*, ainda aqui dando o resultado correto.

No entanto, há pelo menos dois grupos de construções que parecem ser exceções a esse princípio: em ambas um pronome não-reflexivo tem antecedente dentro da mesma oração.

O primeiro caso é ilustrado por

(2) Celinha me deu um retrato dela.

Nessa frase, é perfeitamente possível entender *-ela* como referindo-se a *Celinha*, muito embora ambos os elementos estejam na mesma oração. Esses casos parecem estar limitados a complementos de certos nomes, chamados em inglês **picture nouns** ("substantivos de retrato").

O segundo caso é constituído de frases com objeto direto repetido ("pleonástico"), como

(3) O carro, Celinha o lavou.

Repete-se o mesmo fenômeno: um pronome não-reflexivo tem antecedente dentro da mesma oração.

Esses exemplos mostram que o sistema de regras proposto em 11.1.2. precisa ser aperfeiçoado.

Bibliografia

ALVARENGA, Daniel. 1982. *Sobre interrogativa indireta em português.* Dissertação de Mestrado, UFMG.

ANDRADE, Mário de. 1980. *Modinhas imperiais.* Belo Horizonte, Itatiaia.

BACK, Eurico & MATOS, Geraldo. 1972. *Gramática construtural da língua portuguesa.* São Paulo, FTD.

BARROS, Ev'Ângela B. de. 1992. *Transitividade verbal em português.* Dissertação de Mestrado, UFMG.

——. 1993. Revisão das transitividades, incluindo o complemento do predicado. Manuscrito.

BLOOMFIELD, Leonard. 1942. Outline of ilocano syntax. *Language,* 18. (Reimpresso em HOUSEHOLDER, org., 1972.)

BRANDÃO, Cláudio. 1963. *Sintaxe clássica portuguesa.* Belo Horizonte, Imprensa da UFMG.

CAMARA JR., J. Mattoso. 1964. *Princípios de lingüística geral.* Rio de Janeiro, Acadêmica.

——. 1977. *Dicionário de lingüística e gramática.* Petrópolis, Vozes.

CASTILHO, Ataliba T. de, org. 1990. *Gramática do português falado;* a ordem. Campinas/São Paulo, Unicamp/Fapesp. v. 1.

CEGALLA, Domingos P. 1987. *Novíssima gramática da língua portuguesa.* São Paulo, Nacional.

CHOMSKY, Noam. 1957. *Syntactic structures.* The Hague, Mouton.

——. 1965. *Aspects of the theory of syntax.* Cambridge/Massachusetts, Cambridge University Press/MIT Press.

COLE, Peter & MORGAN, Jerry L., orgs. 1975. *Syntax and semantics 3; speech acts.* New York, Academic Press.

COMRIE, Bernard. 1976. *Aspect.* Cambridge, G. B., Cambridge University Press.

CUNHA, Celso & CINTRA, L. F. Lindley. 1985. *Nova gramática do português contemporâneo.* Rio de Janeiro, Nova Fronteira.

DECAT, Maria Beatriz N. 1978. *Movimento de sintagma nominal interrogado em português.* Dissertação de Mestrado, UFMG.

DE FILIPPO, Francisco. 1974. *Supressão de constituintes coordenados em português.* Dissertação de Mestrado, UFMG.

DUBOIS, Jean & DUBOIS, C. 1971. *Introduction à la lexicographie.* Paris, Larousse.

FERNANDES, Francisco. 1944. *Dicionário de sinônimos e antônimos da língua portuguesa.* Porto Alegre, Globo.

FODOR, Jerry A. & KATZ, Jerrold J., orgs. 1964. *The structure of language.* Englewood Cliffs, N. J., Prentice-Hall.

GREENBAUM, Sidney. 1969. *Studies in english adverbial usage.* London, Longmans.

GRICE, H. Paul. 1975. Logic and conversation. In: COLE, Peter & MORGAN, Jerry L., orgs. *Syntax and semantics 3; speech acts.* New York, Academic Press.

GRUBER, Jeffrey S. 1965. *Studies in lexical relations.* Tese de Doutorado, Massachusetts Institute of Technology.

HANKAMER, Jorge. 1971. *Constraints on deletion in syntax.* Tese de Doutorado, Yale University.

HARRIS, Zellig S. 1957. Co-occurrence and transformation in linguistic structure. *Language,* 33. (Reimpresso em FODOR & KATZ, orgs., 1964.)

HAUY, Amini B. 1983. *Da necessidade de uma gramática-padrão da língua portuguesa.* São Paulo, Ática.

HOUSEHOLDER, Fred W., org. 1972. *Syntactic theory 1*; structuralist. Harmondsworth, Penguin.

HUDDLESTON, Rodney. 1984. *Introduction to the grammar of english.* Cambridge, G. B., Cambridge University Press.

ILARI, Rodolfo & GERALDI, J. Wanderley. 1985. *Semântica.* São Paulo, Ática.

ILARI, Rodolfo et alii. 1990. Considerações sobre a posição dos advérbios. In: CASTILHO, Ataliba T., org. *Gramática do português falado*; a ordem. Campinas/São Paulo, Unicamp/Fapesp. v. 1.

JACKENDOFF, Ray S. 1972. *Semantic interpretation in generative grammar.* Cambridge/Massachusetts, Cambridge University Press/MIT Press.

——. 1983. *Semantics and cognition.* Cambridge/Massachusetts, Cambridge University Press/MIT Press.

KAC, Michael B. 1978. *Corepresentation of grammatical structure.* Minneapolis, University of Minnesota Press.

KURY, Adriano da Gama. 1985. *Novas lições de análise sintática.* São Paulo, Ática.

——. 1972. *Gramática fundamental da língua portuguesa.* São Paulo, Livros Irradiantes.

LEMLE, Miriam. 1984. *Análise sintática*; teoria geral e descrição do português. São Paulo, Ática.

LEVINSON, Stephen C. 1983. *Pragmatics.* Cambridge, G. B., Cambridge University Press.

LEVY, Paulette. 1983. *Las completivas objeto en español.* México, El Colegio de México.

LIMA, C. H. da Rocha. 1964. *Gramática normativa da língua portuguesa.* Rio de Janeiro, José Olympio.

MACAMBIRA, José R. 1982. *A estrutura morfo-sintática do português.* São Paulo, Pioneira.

MACIEL, Maximino. 1918. *Gramática descritiva.* Rio de Janeiro, Francisco Alves.

MARTINET, André. 1979. *Grammaire fonctionnelle du français.* Paris, Crédif/Didier.

MATEUS, Maria H. M. et alii. 1983. *Gramática da língua portuguesa.* Coimbra, Almedina.

MORAES, Euzi R. 1971. *O infinito flexionado em português*; uma análise transformacional. Vitória, Imprensa Oficial.

PALMER, F. R. 1964. Sequence and order. *Monograph Series on Languages and Linguistics,* 17. (Reimpresso em HOUSEHOLDER, org., 1972.)

PERINI, Mário A. 1974. *A grammar of portuguese infinitives.* Dissertação de Doutorado, University of Texas. (Edição brasileira: ——. *Gramática do infinitivo português.* Petrópolis, Vozes, 1977.)

——. 1976. *A gramática gerativa*; introdução ao estudo da sintaxe portuguesa. Belo Horizonte, Vigília.

——. 1985. *Para uma nova gramática do português.* São Paulo, Ática.

——. 1989. *Sintaxe portuguesa*; metodologia e funções. São Paulo, Ática.

—— et alii. *Taxonomia dos elementos nominais do português.* (Relatório do projeto "TENPo", em preparação.)

PONTES, Eunice. 1973. *Verbos auxiliares em português.* Petrópolis, Vozes.

——. 1978. Os determinantes em português. In: ——. *Lingüística e ensino do vernáculo.* Rio de Janeiro, Tempo Brasileiro.

RAPOSO, Eduardo P. 1975. Uma restrição derivacional global sobre o infinitivo em português. *Boletim de Filologia,* XXIV.

ROBINS, R. H. 1964. *General linguistic*; an introductory survey. London, Longmans.

SIMÕES, Anilce M. 1974. *Movimento de quantificadores em português.* Dissertação de Mestrado, Unicamp.

Créditos

Uma gramática como esta é trabalho de síntese e utiliza, direta ou indiretamente, um vasto corpo de pesquisas já realizadas. Por outro lado, há muitos tópicos importantes que não foram adequadamente tratados na literatura, seja em geral, seja em relação com o português; por conseguinte, a elaboração desta *Gramática* também exigiu a realização de pesquisa original.

Não creio que seja possível conceder o devido crédito a todos os trabalhos que influenciaram a análise aqui exposta; tentar fazer isso seria realizar uma anamnese da minha vida de lingüista (algo que dificilmente interessaria o leitor). No entanto, algumas obras influenciaram porções relativamente grandes da análise, e em certos pontos não fiz mais do que traduzir a análise de algum outro pesquisador nos termos que utilizo para a exposição da análise neste livro. A seguir, repasso algumas seções da *Gramática*, mencionando, para cada uma, os trabalhos mais relevantes que serviram de base para a análise adotada. Esses trabalhos são citados, como de costume, por nome do autor e data de publicação.

Não estou incluindo nas menções trabalhos de minha autoria, que naturalmente foram extensamente utilizados como base da exposição; destes, os mais importantes são Perini, 1974 (para a análise do infinitivo pessoal); 1985 (para os aspectos pedagógicos) e 1989 (para a análise sintática); acrescente-se o relatório (em elaboração) do projeto TENPo (Taxonomia dos Elementos Nominais do Português), em que estou trabalhando atualmente com Sigrid Teixeira Fraiha, Regina Stela Bessa Neto e Lúcia Fulgêncio.

Devo avisar o leitor que, na maioria das vezes, não segui estritamente as análises citadas; estas apenas serviram de ponto de partida, muitas vezes precioso, para a elaboração da proposta adotada na *Gramática*. Os autores creditados, portanto, não são responsáveis pela análise final, que é de minha autoria

e, para bem ou para mal, me deve ser atribuída.

Capítulo 3

— A seção 3.1., sobre "A frase e a oração", é em parte derivada da passagem correspondente de Huddleston, 1984.

— A análise do auxiliar e dos predicados complexos (3.2.2.2.) é em parte baseada em Pontes, 1973.

Capítulo 4

— Para as funções suboracionais, utilizei a análise do predeterminante de Simões, 1974. Para o SN, vali-me de Pontes, 1978, e Lemle, 1984.

Capítulo 5

— A análise das interrogativas indiretas (seção 5.3.4.) partiu de Alvarenga, 1982.

Capítulo 6

— A descrição das transitividades verbais (6.2.) segue a proposta de Barros, 1992; 1993.

— Para a forma dos complementos oracionais (6.4.), utilizei Levy, 1983.

Capítulo 7

— A noção de "correspondência" deriva da de "transformação", introduzida por Harris, 1957.

Capítulo 8

— O estudo do posicionamento dos clíticos (8.1.4.) é em parte devedor do trabalho de Mateus et alii, 1983.

Capítulo 9

— A conceituação de "aspecto" foi retirada de Comrie, 1976.

Capítulo 11

— A exposição dos papéis semânticos se baseia principalmente em Jackendoff, 1972, que por sua vez segue Gruber, 1965.

— A análise dos elementos anafóricos (pronomes e elipses) segue Jackendoff, 1972, com muita modificação para se adaptar aos fatos do português.

— Para as elipses em estruturas coordenadas (11.2.5.), utilizei De Filippo, 1977; De Filippo, por sua vez, aplicou ao português a análise proposta por Hankamer, 1971.

Índice remissivo

Os números remetem às seções (numeradas) e aos quadros (precedidos de "Q")

gramática do português no Brasil e nem mesmo resolver todos os problemas das gramáticas normativas que até então conhecemos. No entanto, procura descrever as lacunas de nossos compêndios tradicionais e, sempre que possível, apresentar suas propostas sob um novo enfoque, com mais rigor e de acordo com os estudos recentes.

Mário A. Perini é professor de Lingüística na Universidade Federal de Minas Gerais, em Belo Horizonte. Nos últimos anos, tem se empenhado na produção de materiais para a reformulação do ensino de gramática. Além de artigos em revistas especializadas, já publicou *A gramática gerativa*, *Gramática do infinitivo português*, e, pela Ática, *Para uma nova gramática do português* e *Sintaxe portuguesa: metodologia e funções*.

Esta *Gramática descritiva do português* apresenta uma descrição ampla e minuciosa da estrutura sintática, semântica e léxica do português, assim como uma discussão dos objetivos e princípios metodológicos do estudo da gramática.

O livro se destina sobretudo a professores de língua portuguesa e alunos dos cursos de Letras, mas a sua leitura não pressupõe formação em Lingüística. Com preocupação didática, Mário A. Perini pretende contribuir para a renovação do estudo da gramática portuguesa em nossas escolas.

Neste trabalho, o Autor procura ainda evitar os maiores problemas que vêm sendo apontados em nossa tradição gramatical: falta de coerência interna, inadequação às realidades da língua contemporânea e a má formulação e colocação das normas.